

A STUDY OF THE RELATIONSHIP BETWEEN QUALITY OF PBIS
IMPLEMENTATION, SCHOOL CONTEXT, AND ACADEMIC AND BEHAVIORAL
OUTCOMES

Matthew B. Mier

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in
partial fulfillment of the requirements for the degree of Doctor of Philosophy in the School of
Education (School Psychology).

Chapel Hill
2017

Approved by:

Rune J. Simeonsson

Sandra B. Evarrs

Jill V. Hamm

Steven E. Knotek

Linda H. Mason

©2017
Matthew B. Mier
ALL RIGHTS RESERVED

ABSTRACT

Matthew B. Mier: A Study of the Relationship between Quality of PBIS Implementation,
School Context, and Academic and Behavioral Outcomes
(Under the direction of Rune J. Simeonsson)

Positive Behavioral Interventions and Supports (PBIS) is a framework that has been employed in schools in order to promote not only improved behavioral outcomes but academic outcomes as well. PBIS has three core elements: use of a three-tiered prevention model, application of evidence-based practices, and creation of systems to support and sustain implementation. To promote quality implementation of PBIS, the North Carolina Department of Public Instruction developed a three-level system of school recognition based upon level of training attained, fidelity of implementation, and whether schools demonstrated positive trends in data.

Schools in each of the three recognition levels were included in this study to gain information about the relationship between PBIS quality, school contextual variables, suspension rates, and percentage of students demonstrating proficiency on End-of-Grade Reading and Math Assessments. Contextual variables of interest were School Economic Status and Setting. Each school included in the study was identified as being a school of Low or Moderate/High Economic Status and as a City, Suburb, Town, or Rural school. Three research questions were posed pertaining to the relationship between PBIS quality, Economic Status, suspension rates, and EOG Reading and Math performance. Two-way ANOVAs were used to test the hypotheses that there would be an interaction between PBIS Level and Economic Status on each of the three dependent variables. No significant

interactions were found. There was a significant effect of Economic Status on suspension rate. There were also significant effects of PBIS Level and Economic Status on EOG Reading and Math Assessment performance.

Additionally, an exploratory analysis was conducted to examine possible interactions between PBIS Level, Economic Status, and Setting. Due to the small number of schools in some groups, Setting was condensed into two groups: City/Suburb and Town/Rural. Three-way ANOVAs were used to explore effects of the three independent variables on suspension rate and EOG Reading and Math performance. There was a significant interaction between PBIS Level and Setting and a significant main effect of Economic Status on suspension rate. With regard to both Reading and Math achievement, there was a significant interaction between Economic Status and Setting and a significant main effect of PBIS Level.

To my sons Holden and William, who are constant sources of
encouragement and inspiration.

ACKNOWLEDGEMENTS

I have been fortunate to receive guidance and support from a number of individuals during the course of my dissertation work. I would first like to express my sincere gratitude to my dissertation chair, Dr. Rune Simeonsson, and my committee members, Dr. Sandra Evarrs, Dr. Jill Hamm, Dr. Steve Knotek, and Dr. Linda Mason, for their willingness to give of their time and expertise throughout the dissertation process. I would also like to thank Dr. William Ware for providing guidance regarding my statistical analyses. I am especially grateful to my wife Sarah for supporting and believing in me throughout this process and for her role as an amazing mother to our two children. I could not have gotten to where I am without her. To my mother Jeannie, I want to express thanks for instilling in me the traits of positivity and persistence, and for her willingness to provide support in any number of ways throughout my life and during pursuit of my doctorate. I would also like to acknowledge other members of my family: Bob, Dianne, Mike, Hillary, Shannon, Sarah, Ollie, and Orion, all who helped and encouraged me the past four years. Finally, I have been shaped as a student and professional by a number of people, including but certainly not limited to Dr. Rodney Clark, Ryan Delaney, Cecelia Breazeale, Marc Giandenoto, Dr. Cathy Metcalf-Shaw, and Dr. Lynne Taylor. I appreciate the time that they invested in helping me to develop my knowledge and skills.

TABLE OF CONTENTS

LIST OF TABLES	x
LIST OF FIGURES	xii
CHAPTER 1: INTRODUCTION	1
CHAPTER 2: REVIEW OF THE LITERATURE	6
Characteristics of PBIS	6
The Relationship between Behavior, School Climate, and Academic Outcomes	10
Impact of PBIS Implementation on Academic and Behavioral Outcomes on a Small Scale.....	13
Fidelity and Large-Scale PBIS Implementation	15
Role of School Context.....	24
PBIS Implementation in North Carolina.....	34
CHAPTER 3: RATIONALE FOR STUDY AND RESEARCH QUESTIONS	40
CHAPTER 4: METHODOLOGY	44
Selection of Schools.....	44
Procedure	44
PBIS Level	45
Measures Used for Identification of PBIS Recognition Level	46
School-Wide Evaluation Tool.....	46
Benchmarks of Quality	47
Implementation Inventory Online	47

School Economic Status	48
Setting	48
Outcome Measures.....	49
Suspension Rates	49
End-of-Grade Assessments	49
Analysis.....	50
CHAPTER 5: RESULTS	52
Research Question 1	52
Tests of Homogeneity and Normality	54
Hypothesis 1.....	55
Research Question 2	56
Hypothesis 2.....	57
Research Question 3	59
Hypothesis 3.....	60
Outcomes among Schools Grouped by PBIS Level, Economic Status, and Setting.....	62
Suspension Rates	62
Reading Proficiency.....	66
Math Proficiency.....	70
CHAPTER 6: DISCUSSION.....	74
Overview	74
Study Findings	75
Suspension Rates	75
Reading Proficiency.....	77

Math Proficiency.....	80
School Setting	83
Limitations	85
Implications and Future Directions.....	87
REFERENCES	91

LIST OF TABLES

Table 1: Number of school settings by group.....	52
Table 2: Short-term suspension rates per 100 students by economic status and PBIS level.....	53
Table 3: Short-term suspension rates per 100 students by economic status and PBIS level with adjusted outliers	55
Table 4: Analysis of Variance: Short-term suspension rates per 100 students by economic status and PBIS level with adjusted outliers	56
Table 5: Percentage of students who passed the EOG Reading Assessment	57
Table 6: Analysis of Variance: Percentage of students who passed the EOG Reading Assessment	58
Table 7: Tukey's HSD: PBIS Level and percentage of students who passed the EOG Reading Assessment.....	59
Table 8: Percentage of students who passed the EOG math assessment.....	60
Table 9: Analysis of Variance: Percentage of students who passed the EOG Math Assessment	61
Table 10: Tukey's HSD: PBIS Recognition Status and percentage of students who passed the EOG Math Assessment.....	61
Table 11: Suspension Rate per 100 students among schools grouped by PBIS Level, Setting, and Economic Status	63
Table 12: Analysis of Variance: Short-term suspension rates per 100 students by PBIS level, economic status, and setting with adjusted outliers.....	65
Table 13: Simple main effects of PBIS Level and Setting on Suspension Rates	66
Table 14: Percentage of students who passed the EOG Reading Assessment among schools grouped by PBIS Level, Setting, and Economic Status	67
Table 15: Analysis of Variance: Percentage of students who passed the EOG Reading Assessment by PBIS Level, Setting, and Economic Status	68
Table 16: Simple main effects of Economic Status and Setting on EOG Reading Assessment Proficiency.....	70

Table 17: Percentage of students who passed the EOG Math Assessment among schools grouped by PBIS Level, Setting, and Economic Status	70
Table 18: Percentage of students who passed the EOG Math Assessment among schools grouped by PBIS Level, Setting, and Economic Status	72
Table 19: Simple main effects of Economic Status and Setting on EOG Math Assessment Proficiency.....	73

LIST OF FIGURES

Graph 1: Interaction between PBIS Level and Setting on Suspension Rate.....	66
Graph 2: Interaction between school setting and economic status on reading proficiency ..	69
Graph 3: Interaction between Setting and Economic Status on math proficiency	73

CHAPTER 1: INTRODUCTION

At the turn of the 21st century, the North Carolina Department of Public Instruction was faced with troubling data regarding student behavior and discipline in schools across the state. Out-of-school suspensions caused students to miss over one million instructional days in the 2001-2002 school year (Irwin & Algozzine, 2005, p. 2). There had been a 27% increase in long-term suspensions over the previous year, and over half of students receiving these suspensions were black or multi-racial. Additionally, students who received suspensions were less likely to demonstrate proficiency on standardized achievement tests (Irwin & Algozzine, 2005, p. 3). To address needs in this area, state law required districts to collaborate with communities in the development of plans to foster secure and respectful schools (Irwin & Algozzine, 2005, p. 2). In search of a framework that would assist in addressing discipline challenges and promoting environments that allowed students to reach their potential, North Carolina turned to Positive Behavioral Interventions and Supports (PBIS).

PBIS, alternately referred to as Positive Behavior Support (PBS) or School-Wide Positive Behavior Support (SWPBS), is a framework that applies the principles of Response to Intervention to behavior. PBIS employs a tiered approach to intervention that is preventive in nature and emphasizes analysis of data, evidence-based and culturally-responsive practices, and creation of structures that support sustainability and fidelity (Technical Assistance Center on Positive Behavioral Interventions and Supports, 2010). In contrast to the punishment-focused zero-tolerance discipline approaches that were popular in

the 1990's, PBIS emphasizes teaching and reinforcement of desired behaviors and utilizes data analysis to identify functions of problematic behavior and assist in identifying replacement behaviors. Studies at the school and district level have supported a relationship between PBIS implementation and improved academic and behavioral outcomes: Luiselli, Putnam, Handle, and Feinberg (2005) found improvements in performance on reading and math assessments and lower rates of office discipline referrals (ODRs) following PBIS implementation at an elementary school. Lassen, Steel, and Sailor (2006) studied PBIS implementation at a middle school, finding a decrease in ODR and suspension rates and increased math scores. At the district level, McIntosh, Bennet, and Price (2011) found an overall decrease in ODRs following PBIS implementation, and a higher percentage of students meeting reading and math standards among schools implementing PBIS with a moderate to high degree of fidelity.

Initial implementation of PBIS in North Carolina, which began at Oak Grove Elementary School in Durham County during the 2001-2002 school year, also produced promising results (Irwin & Algozzine, 2005, p. 1). Oak Grove saw a reduction in suspension rates, and other early implementing schools saw decreases in ODRs as well (Irwin & Algozzine, 2005, p. 12, 13). As a result, administrators and teachers gained time that would have been spent on the discipline process, freeing up teachers to spend more time teaching their students. With the reduction of suspension rates, students also spent more time in the classroom. The initial implementation of PBIS in North Carolina was considered successful, and additional resources were allocated to expand application of the framework in schools across the state. Over the next 12 years, PBIS implementation in North Carolina grew

rapidly: by the 2012-2013 school year there were 1102 participating schools across the state (Public Schools of North Carolina, 2014).

North Carolina was not alone in embracing PBIS as a pathway to improved academic and behavioral outcomes. According to pbis.org, almost 14,000 schools across the country were implementing PBIS by 2013 (Reinke, Herman, & Stormont, 2013). Studies that have examined large-scale implementation of PBIS have generally found a relationship with improved behavioral and academic outcomes. Muscott, Mann, and LeBrun (2008) found reductions in ODRs and suspensions as well as improved math scores in PBIS schools in New Hampshire. Bradshaw, Mitchell, and Leaf (2010) also found a reduction in suspension and ODR rates among schools implementing PBIS. Horner et al. (2009) studied outcomes among 53 schools in Illinois and Hawaii, finding increases in perceptions of school safety and higher scores on reading assessments following PBIS implementation.

Of importance, these studies also examined the extent to which schools were able to implement PBIS with fidelity, finding that many were able to achieve high-fidelity implementation. In recent years there has been an increased research focus on the intricate process of large-scale implementation. While many factors impact the degree to which scaling up is successful, fidelity of implementation and whether interventions are adaptable to the diverse needs of a variety of schools are two critical variables. Of the four core elements identified by the Technical Assistance Center on Positive Behavioral Interventions and Supports as necessary for large-scale implementation, two were supports for accurate practice implementation and ongoing analysis to ensure effectiveness across contexts (Technical Assistance Center on Positive Behavioral Interventions and Supports, 2010, p. 35).

As PBIS was scaled up throughout North Carolina, the state's PBIS Initiative took steps to ensure that both fidelity and adaptation to context were addressed. Contextually-relevant practices were promoted during training, which addressed the importance of considering the unique culture and environment of each school when implementing PBIS (Public Schools of North Carolina, n.d.2). Multiple layers of support were provided to districts and schools to promote implementation fidelity in the form of training, coaching, and technical assistance (Reynolds, Irwin, & Algozzine, 2009). Additionally, a multi-tiered system of recognition of schools was created to encourage and reinforce completion of training, implementation with fidelity, creation of school-level systems to support PBIS, and ultimately improvements in behavioral and academic outcomes. The number of schools recognized for quality implementation grew from 68 in the 2007-2008 school year to 513 in the 2013-2014 school year (Public Schools of North Carolina, 2014, "PBIS Green Ribbon Schools 2013-2014", n.d., "PBIS Model Schools 2013-2014", n.d., "PBIS Exemplar Schools 2013-2014", n.d.). This number represented a substantial portion of the total number of schools in the state: 21% of schools earned one of the three levels of recognition.

The recognition system developed in North Carolina provides an opportunity to analyze implementation in the state and to contribute to the research base regarding scaling up of the PBIS model. While previous research has generally supported PBIS as a means of supporting student academic and behavioral success, additional information is needed. Many of the studies that examine PBIS outcomes have limitations, including small sample sizes and failure to apply tests of significance (Luiselli et al., 2005; McIntosh et al., 2011, Muscott et al., 2008). Additionally, most studies used only one indicator of quality of implementation, generally in the form of score on a fidelity measurement tool. Research by Pas and

Bradshaw (2012) found that different measures of fidelity varied in their relationships to academic and behavioral outcomes, supporting the need for a more robust way of examining quality of implementation. Finally, very little research has been conducted on the role that context plays in large-scale PBIS implementation. The research that has been done has found a relationship between contextual variables and fidelity of implementation at the school and classroom level (Pas & Bradshaw, 2012, Pas, Waasdorp, and Bradshaw, 2014).

The purpose of this study is to examine behavioral and academic outcomes in a diverse set of schools across North Carolina to determine if identification of quality PBIS implementation using a multifaceted assessment model is associated with improved behavioral and academic outcomes. PBIS was identified by the North Carolina Department of Public Instruction as a “way to impact the learning environment in all schools in order to support high student performance and to reduce behavior problems” (Irwin and Algozzine, 2008, p. 2). Specifically, this study seeks to investigate whether PBIS has been effective in these areas by analyzing whether recognition of varying levels of quality of implementation is associated with higher End of Grade assessment scores and lower rates of suspension. Additionally, an important consideration of large-scale implementation is whether outcomes are consistent across a variety of contexts. This study will examine the contextual variables of school economic status and school setting to determine whether these variables moderate the effectiveness of PBIS. Investigation of the relationship between fidelity of implementation and contextual variables and behavioral and academic outcomes will provide information about the effectiveness of PBIS implementation in diverse schools across North Carolina and will contribute to the research base regarding factors impacting large-scale implementation.

CHAPTER 2: REVIEW OF THE LITERATURE

Characteristics of PBIS

Although PBIS was originally developed as a result of research into effective school-based interventions for students with behavior disorders in the 1980's, investigation into its potential as a school-wide framework began in the late 1990's (Sugai & Simonsen, 2012). At the time, the wisdom and effectiveness of zero-tolerance discipline policies was beginning to be questioned (Skiba & Peterson, 1999, 2000; Morrison & D'Incau, 1997). Researchers at the University of Oregon found that there was a need for development of methods focusing on prevention, use of evidence-based interventions, and development of systems-level supports. This need was codified into law in 1997 with the passage of amendments to the Individuals with Disabilities Education Act, which required school-based teams to consider positive behavioral intervention strategies and to conduct functional behavior assessments and develop behavioral intervention plans as part of the disciplinary process for students with disabilities (Sugai et al., 2000). Since that time, PBIS has developed into a widely used whole-school approach to promoting improved behavioral and academic outcomes.

Many diverse approaches to school discipline and behavior management have similar goals: to promote prosocial behavior, decrease antisocial behavior, and create safe environments that foster learning. PBIS and zero-tolerance approaches share roots in behavioral psychology. They differ, however, in the extent to which they each rely on reinforcement and punishment. Zero-tolerance policies establish rigid consequences that theoretically punish inappropriate behavior, making future misbehavior less likely. In reality,

however, studies have found that zero-tolerance policies do not reduce future misbehavior or create a healthier school climate (Sugai & Horner, 2002, American Psychological Association Zero Tolerance Task Force, 2008). PBIS separates itself from zero-tolerance policies in part by emphasizing a focus on teaching and reinforcing prosocial behavior rather than on consequences.

PBIS draws upon the theory that human behavior is learned, and is maintained through consequences in the environment (Sugai et al., 2000). At times, these learned behaviors may be maintained by teachers unintentionally, as when a teacher sends a student to the office for exhibiting disruptive task-avoidant behavior. Application of PBIS principles involves identifying and explicitly teaching expected behaviors in the classroom, cafeteria, playground, and all other areas of the school. Exhibition of the target behaviors can be reinforced through many contingencies, including class-wide reinforcement, verbal praise, or the use of a token system. For some students who engage in inappropriate behavior, it may be necessary to conduct an analysis of the function of that behavior and to develop individual plans to teach new, more adaptive behaviors. While students may receive consequences for misbehaving at PBIS schools, there is a focus on teaching and rewarding appropriate behavior rather than punishing inappropriate behavior. Promotion of a healthy school climate occurs at PBIS schools not through a prescribed set of interventions but rather through application of a framework with three core elements: prevention, evidence-based practice, and systems implementation (Sugai & Horner, 2006).

To reduce the number of students who exhibit significant difficulty functioning in the school environment, PBIS employs a public health model of prevention. This model, which was first implemented in a school setting in the 1950's, incorporates a three-tiered approach

to prevention (Merrell, & Buchanan, 2006). The focus of the primary level is on prevention of problem behaviors through provision of universal supports. These supports may include teaching and reinforcing classroom and schoolwide expectations, increased monitoring in “trouble areas” around the school, and promotion of prosocial behavior through selection of curricula to address areas of need, such as a school-wide bullying prevention program or social skills curriculum. An effective universal support system is expected to support about 80% of students’ needs.

For the approximately 20% of students who exhibit problem behaviors in spite of quality universal supports, secondary interventions may be necessary (Technical Assistance Center on Positive Behavioral Interventions and Supports, 2010, p. 20). At the secondary level, a greater level of support is provided to students demonstrating need. This can include skill development through small group instruction, use of a point sheet to provide reinforcement for desired behaviors, or use of a Check-In/Check-Out procedure in which students review expectations in the morning and discuss performance at the end of the day. Use of secondary interventions is generally brief in nature, and phased out when students have developed skills or demonstrated behaviors that allow for success with universal supports.

About 5% of students may not respond to primary support or secondary interventions. For these students, intensive, individualized support is provided at the tertiary level. Students who require this level of support typically present with fairly severe behavioral problems that may include aggressive or oppositional behavior. Often, tertiary support includes a functional behavior assessment to gain information about the contingencies maintaining the behaviors of concern. Tertiary intervention may focus on identification, teaching, and

reinforcement of alternative, appropriate behaviors. Typically, behavior is monitored more closely than at the primary or secondary levels, and plans are reviewed frequently to determine their effectiveness.

The second core element of PBIS is use of evidence-based practices. At each level of the continuum, there is a focus on application of interventions supported by research findings, which serves to reduce wasted time, effort, and money, allowing for efficient and effective services for students. Selection of interventions and practices are guided by their research base, and priority is given to interventions that have been validated by multiple studies (Sugai & Horner, 2006; Technical Assistance Center on Positive Behavioral Interventions and Supports, 2010, p. 15). In situations in which there is a lack of evidence-based interventions, new interventions are implemented initially on a small scale, and outcomes are evaluated frequently to monitor effectiveness (Sugai & Horner, 2006).

The final core element of PBIS is a focus on creation of systems to support and sustain implementation (Sugai & Horner, 2006; Sugai & Simonsen, 2012; Technical Assistance Center on Positive Behavioral Interventions and Supports, 2010, p. 15). This includes many macro- and school- level factors identified by Domitrovich, et al. (2008) as impactful of the quality of implementation, including creation of policy, identification of sources of funding, provision of quality training, and access to ongoing coaching and support. Additionally, successful implementation of the PBIS framework requires identification of school-level needs and creation of corresponding measurable goals to address those needs (Domitrovich, et al., 2008). Analysis of data is necessary at the school, district, and state level. School-level analysis can identify degree of fidelity of implementation, effectiveness of each of the three tiers of prevention, and whether

interventions are associated with improved outcomes. District- and state-level analysis of data is necessary to determine overall effectiveness of initiatives and to identify areas of need that may need to be addressed through training or increased support. The ultimate goal of PBIS is not only improved behavioral outcomes but improved academic performance of students as well. Evidence that supports the relationship between academic and behavioral outcomes as well as the effectiveness of PBIS at improving these outcomes will be reviewed below.

The Relationship between Behavior, School Climate, and Academic Outcomes

As noted above, PBIS was selected for implementation in North Carolina as a means of promoting not only improved behavior but also higher student achievement. There is a large base of evidence to support the existence of a relationship between behavior, school climate, and academic performance. Glew et al. (2005) found that bullying victimhood was associated with lower academic achievement and feeling unsafe. Students who played the dual roles of bully and victim were also more likely to have lower academic achievement. Gietz and McIntosh (2014) found that, when controlling for the effects of district variables as well as school socioeconomic (SES) level, student perceptions of victimization at school, feeling safe at school, and knowledge of school expectations were predictors of academic achievement at multiple grade levels. Student perceptions accounted for between 6 and 9% of variance in academic scores.

Studies have also linked behavior and academic performance over time. Miles and Stipek (2006) examined the relationship between social skills and literacy development among elementary students, hypothesizing that low literacy skills would predict higher rates of aggression. Data collected over a four-year period indicated that first grade literacy skills

predicted aggressive behavior in third grade, and that literacy skills in third grade also predicted aggression two years later. There was some evidence that behavior predicted academic outcomes as well: prosocial behavior in first grade predicted literacy skills in third grade. Algozzine, Wang, and Violette (2010) also found a relationship between academic and behavioral outcomes among elementary students. A negative relationship was found between teacher-assessed problem behaviors and academic competence, as well as between low social skills and academic achievement.

Darney, Reinke, Herman, Stormont, and Ialongo (2013) found that behavioral and academic problems in first grade predicted even more distal outcomes. In a study of 574 students, data were collected on the extent to which first graders exhibited aggressive behavior, oppositional behavior, and attention problems. Data were also collected on the students' performance on standardized reading and math assessments. Students with behavior problems in the first grade were more likely to be the recipient of special education services in the twelfth grade, and were more likely to have a diagnosis of conduct disorder. Students with low academic scores in first grade were more likely to have continued low academic performance in the twelfth grade, and were less likely to graduate from high school. As might be expected, though, students with the greatest risk of future difficulties were those who presented with co-occurring academic and behavioral problems. Problems in both of these domains were associated with poor reading and math scores, special education placement, failure to graduate, and increased use of mental health services.

Although there is disagreement regarding whether academic skills and behavior are causally linked, research consistently supports the existence of a relationship. A common conclusion by researchers in this area is that early intervention and quality practices in both

academics and behavior management are critical (Algozzine et al, 2010; Darney et al, 2013; Miles & Stipek, 2006). With its emphasis on prevention and evidence-based practices, PBIS is positioned to contribute to improvements in behavioral and academic outcomes through several mechanisms. PBIS may decrease incidence of bullying and increase safe school environments through universal promotion of prosocial skills and secondary and tertiary support for the small percentage of students with increased behavioral needs. Research has found positive effects of a classroom environment that provides supports similar to those found in a PBIS framework. Hamre and Pianta (2005) found that at-risk students who were placed in classrooms with high levels of emotional support attained similar academic achievement scores as their low-risk peers, in comparison to at-risk students in low- or moderate-support classrooms, who had lower scores.

Additionally, a focus on use of reinforcement of appropriate behavior rather than use of exclusionary punishment may result in a decrease in students' lost instructional time as a result of suspensions. Teachers may also increase their instructional time as classroom management becomes more efficient and they have to dedicate less time to disciplinary issues. Additionally, in instances in which problem behavior is sustained through avoidance of academic tasks, PBIS may be beneficial. The use of data-based behavior assessment to determine antecedent events and consequences of behavior may help to identify associations between problem behavior and academic areas and point to remediation as part of a plan to address the behavior.

Impact of PBIS Implementation on Academic and Behavioral Outcomes on a Small Scale

Research on behavioral and academic outcomes of PBIS implementation has been conducted at the school, district, and state levels. Typical methods of measuring behavioral outcomes have included examining the rate of suspensions or office discipline referrals (ODRs) over time or between schools implementing PBIS with varying degrees of fidelity. Frequently, reading and math scores on summative evaluations have been used to assess the effects of PBIS on academic outcomes. Studies have typically found reductions in problem behaviors. The evidence with regard to academic outcomes has been mixed: although evidence has suggested some improvements in areas of academic functioning, limitations in research or statistical methods prevent firm conclusions from being drawn.

Lassen, Steele, and Sailor (2006) and Luiselli, Putnam, Handler, and Feinberg (2005) examined behavioral and academic outcomes of PBIS implementation in two urban schools. The participants in Luiselli et al.'s (2005) study attended an elementary school in the mid-west United States. The school's leadership implemented PBIS in an attempt to address behavioral and academic problems as well as low staff morale. Behavioral outcomes of interest were the rate of ODRs and suspensions, and academic performance was gauged through examination of student performance in reading and math standardized assessments.

Results indicated that after an initial rise in the rate of ODRs during the first three months of implementation there was a decrease that was sustained throughout the remainder of the study. With regard to suspensions, rates were typically lower than pre-intervention levels during the first half of the second and third year of the study. During the second half of these years, however, they increased to levels close to or higher than those observed pre-

intervention. To assess academic performance, the authors examined average percentile ranks in the areas of reading comprehension and math. Large increases were seen in each of these areas from pre- to post-intervention (18% increase in reading comprehension, 25% increase in math). Results obtained have limited generalizability, however. Tests of significance were not performed, so while there were observed differences in ODR rates and academic assessment performance, it is not known whether these differences were statistically significant.

Lassen, Steel, and Sailor (2006) also assessed the relationship of PBIS implementation and student behavior and academic performance through examination of ODRs, suspension rate, and performance on standardized reading and math assessments. The participating school was an urban middle school in the mid-west United States. Staff at the school were provided with PBIS training the middle of the first year of the study, during which time universal supports were introduced. Training continued throughout the study, and by the third year secondary supports were in place.

The rate of ODRs per student decreased significantly over the course of the study, as did the rate of suspensions per student. Academic data presented a mixed picture. Significant gains were not obtained in the area of reading, but math scores increased significantly from baseline to years 2 and 3. Regression analyses of the relationship between behavior and academic performance indicated that while the number of ODRs and suspensions a student received significantly predicted academic achievement, the effect sizes were small, accounting for only 1 to 2% of test score variability.

Fidelity and Large-Scale PBIS Implementation

While the previously mentioned studies provided evidence to support the hypothesis that use of PBIS is associated with improved behavioral or academic outcomes, the small scale of each limits the ability to generalize their findings to large-scale application of the PBIS model in diverse settings. A common stumbling block to scaling up of intervention programs noted by Hoagwood and Johnson (2003) is that much of the development and assessment of evidence-based practices takes place in highly controlled environments under circumstances that may not translate to real-world settings. School settings were utilized by Luiselli et al. (2005) and Lassen et al. (2006), but consultation and support was provided by either out-of-state doctoral-level psychologists or the researchers themselves. The level of support and expertise provided to participants may have exceeded that which might realistically be expected in large-scale efforts. Additionally, each study analyzed data from only one school and performed limited analyses regarding the relationship between fidelity of implementation and outcomes.

In reality, the process of scaling up interventions from pilot testing to large-scale application is an intricate process: fidelity of implementation can be difficult to obtain but is important to outcomes. Hoagwood and Johnson (2003) described a two-way adaptation that often takes place when evidence-based practices are put into place: changes are made to the school setting in order to accommodate the intervention, and, importantly, adaptations of the intervention occur in order to meet practical considerations. Similarly, Rohrbach, Grana, Sussman, and Valente (2006) noted in their review of research on intervention translation that programs are often modified by schools, potentially impacting their effectiveness. Quality implementation of interventions is dependent on a number of variables, including

support from administration and staff, an infrastructure to provide consultation and training, and development of internal mechanisms of support and evaluation (McDougal, Sheila, & Martens, 2000). In light of the importance of fidelity and ongoing assessment of implementation, multiple evaluation tools have been developed for PBIS. One such tool, the Schoolwide Evaluation Tool (SET) has been frequently utilized by researchers studying PBIS as well as by districts and states, including North Carolina, which have implemented the framework (Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. B., 2004).

The SET provides information about the degree to which seven critical features are present in schools implementing PBIS. In keeping with the structural aspects identified by McDougal et al. (2000), the SET assesses the extent to which there is active administrator support, involvement and provision of district support through policy development, training opportunities, and data collection. Additionally, the SET provides information about how well schools incorporate the behavioral elements of PBIS: definition of school-wide behavioral expectations, teaching of expectations, provision of rewards for following expectations, consistent implementation of a continuum of consequences for problem behavior, monitoring of problem behavior patterns and data-based decision-making. Assessment of these areas is conducted by an outside observer, such as a district- or state-level PBIS consultant, who conducts school observations and interviews and reviews data, reports, and other PBIS-related materials. While various criteria are used by schools and researchers to indicate whether fidelity has been obtained, an oft-cited standard is attainment of 80% of the criteria in the area of teaching expectations and 80% of total criteria (referred to as the “80/80 criteria”).

A number of studies that have examined PBIS implementation on a district- or state-wide scale have found that fidelity of implementation may significantly impact outcomes. McIntosh, Bennett, and Price (2011) investigated fidelity of implementation and academic and behavioral outcomes in schools in an urban Canadian school district that had been implementing PBIS for as long as five years. PBIS fidelity was assessed using the total SET score. Academic outcomes of interest were the percentage of students meeting or exceeding standards in reading and math as assessed on standardized tests, and the rate of ODRs per 100 students was used to assess behavioral outcomes. SET scores indicated that only two of nine schools assessed were implementing PBIS with high fidelity (at least 80% of total criteria); four were implementing with moderate fidelity (70-79% of total criteria), and three were implementing with low fidelity (less than 70% of total criteria).

The authors found that schools that were implementing PBIS with moderate to high fidelity had higher percentages of students who met or exceeded standards in the areas of reading and math. Additionally, moderate- and high-fidelity schools had levels of academic achievement that were higher than the district average, despite the fact that these schools had a higher percentage of students from low-income families than the district average. This finding supports the notion that PBIS implementation is related to higher academic achievement, and indicates that the relationship may be moderated by fidelity of implementation. The authors also found that PBIS schools saw a drop in the rate of ODRs from the first year of the study to the second. Several factors limit the conclusions that can be drawn from the study. The authors did not provide information about the outcomes associated with the varying levels of fidelity. In addition, they did not test the significance of observed differences.

Horner et al. (2009) sought to study PBIS implementation on a larger scale, collecting data over the course of four years on 53 elementary schools in Hawaii and Illinois. Thirty of the schools implemented PBIS initially, while the other schools served as a control/delay group. The researchers provided training to state PBIS personnel, but provided no direct support to schools. Variables of interest, which were measured multiple times over the course of the study, included fidelity as assessed by the SET, perceived student safety, ODR rates, and the proportion of third graders meeting or exceeding state reading standards, as assessed on state summative tests.

With regard to fidelity of implementation, a time x condition analysis showed significant effects on total SET scores between groups and from pre- to post-intervention. A second random coefficients analysis indicated significant differences between total SET scores before and following training for the Treatment group, followed by non-significant differences in subsequent years. The Control/Delay group, on the other hand, did not have significantly different pre-and post-training SET scores, but had significant increases in the following years. As a possible explanation, the authors pointed to decreases in the amount of training provided to school personnel in the Hawaiian Control/Delay group. This example of the real-world challenges of scaling up interventions may explain the differences in average total SET scores among the Control/Delay schools in Illinois and Chicago (81% and 79%) as compared to the Hawaiian schools (59%).

Perceptions of school safety significantly increased in both groups following PBIS implementation, providing evidence that PBIS implementation is associated with positive changes to the school environment. While the rate of ODRs following PBIS implementation was lower than the national average among schools utilizing SWIS, a behavior-tracking

database, no pre-intervention data was collected, limiting the ability to identify a relationship. Scores on state reading assessments were significantly higher following PBIS implementation, though the authors stressed the need for further research to explore the relationship between PBIS and academic outcomes.

Simonsen et al. (2012) also studied the effects of PBIS implementation in Illinois, collecting data on a total of 382 schools (274 elementary, 91 middle, and 17 high schools). The SET was used to examine differences in behavioral and academic outcomes among low- and high-fidelity schools. The 80/80 standard was used by the authors to designate high-fidelity status. Behavioral outcomes that were examined were ODR rates, number of suspensions, and out-of-school suspension days, while the percentages of students meeting or exceeding standards on math and reading assessments were included to examine academic outcomes.

Hierarchical linear modeling indicated that fidelity of implementation did not contribute significantly to ODR rates. Fidelity did contribute significantly, however, to out-of-school suspension days and the rate of suspensions. Schools that implemented with fidelity had significantly lower rates of out-of-school suspension days and total suspensions. Regarding academic outcomes, fidelity did not significantly contribute to the reading model, but significant differences existed between the low- and high-fidelity schools in the area of math (80.63% of students in high-fidelity schools met or exceeded math standards, compared to 72.89% of students in low-fidelity schools). While causality cannot be inferred, the results obtained by Simonsen et al. (2012) add to the evidence base linking PBIS implementation with improved behavioral and academic outcomes. The results also further support the

notion that fidelity impacts the relationship between PBIS and academic and behavioral outcomes.

Muscott, Mann, and LeBrun (2008) provided additional evidence of a relationship with their examination of PBIS scale-up efforts in New Hampshire. Data were collected on academic and behavioral outcomes among 29 schools that were among the first in the state to begin implementation. The authors sought to determine the ability of the state's training and assistance programs to support fidelity of implementation, as well as whether implementation was associated with decreases in problem behavior, increased availability of students for instruction, and improved academic functioning. The researchers monitored fidelity over time with the SET, finding that 15 of the 28 schools met the 80/80 criteria within three months of implementation. Twenty one of 24 participating schools met the criteria by fall of the second year, and fidelity was maintained over time among almost all schools who reached the 80/80 criteria.

Behavior was assessed by examining ODR data entered into the SWIS database. Fourteen of 20 schools reached their identified goals regarding the percentage of students with one or no ODR, and the 22 schools that reported data reduced their total number of ODRs by 28% between years 1 and 2 of the study. While a reduction was observed, it was noted that the average rate of ODRs at these schools remained higher than the SWIS national average. The number of suspension was also decreased over time, with in-school suspensions decreasing by 31% and out-of-school suspensions decreasing by 19%. These reductions were primarily seen at the middle and high school levels. The authors noted, though, that hundreds of hours of instructional time were saved at the elementary level when reductions in suspensions and ODRs were combined.

A review of academic indicators found positive outcomes in the area of math. Overall, 73% of schools that met the 80/80 criteria saw improved math scores. Eleven of 13 elementary or multilevel schools saw increased percentages of third grade students who met or exceeded standards. On the third grade reading assessment, however, less than a third saw improvement. Unfortunately, the authors did not conduct analyses to assess the significance of differences of between- or within-group differences, limiting the conclusions that can be drawn.

Finally, much research has taken place regarding the effectiveness of PBIS implementation in the state of Maryland. Bradshaw, Mitchell, and Leaf (2010) conducted a randomized controlled effectiveness trial, assigning a total of 37 schools to either a PBIS treatment or control group. Data were collected on fidelity of implementation through the use of the SET as well as the Effective Behavior Support (EBS) survey, a measure that is completed by staff and provides information about the extent to which PBIS systems are in place. Analysis of fidelity data using repeated-measures general linear models indicated significant intervention effects as measured by both the SET and EBS. The authors noted that all schools trained in PBIS were able to attain high-fidelity implementation, and sustained this level of implementation throughout the study.

Behavioral outcomes of interest were the rates of ODRs and suspensions, which were entered by participating schools into the SWIS database. A significant reduction in both of these areas was observed in PBIS schools: the percentage of students receiving a major or minor ODR as well as the number of ODRs per students decreased significantly throughout the study. PBIS schools also saw a significant decrease in the percentage of students who received suspensions. With regard to academic outcomes, measured by performance on

third- and fifth-grade reading and math assessments, possible differences between the PBIS and control schools were analyzed using t tests: while the authors noted that in general greater gains were seen in the PBIS group, no significant differences were found in any of the subjects or grade levels.

Implementation on a larger scale was analyzed by Pas and Bradshaw (2012), who examined the relationship between PBIS fidelity and behavioral and academic outcomes in 421 elementary and middle schools in Maryland. The authors utilized three measures of fidelity: the SET, the Implementation Phases Inventory (IPI), which was completed by schools' PBIS coaches and provided information about the degree to which features connected to four stages of implementation were in place, and the Benchmarks of Quality (BoQ), which assessed fidelity in ten areas of implementation and was completed by members of the school PBIS team as well as the school's coach. Behavioral outcomes were truancy and suspension rates, and academic outcomes of interest were math and reading scores on standardized assessments.

The authors utilized a structural equation model in order to analyze the relationship between fidelity of implementation and outcomes while adjusting for a number of school-level variables, including the number of years since PBIS training, school size, student/teacher ratio, percentage of certified teachers, and student mobility. When these variables, as well as baseline academic and behavioral performance, were controlled for, IPI scores were significantly related to math and reading scores as well as the behavioral indicator of truancy: higher fidelity was associated with improved academic scores and lower truancy rates. Of interest was the finding that neither SET nor BoQ scores were significantly

related to behavioral or academic outcomes. With regard to the SET, the authors hypothesized that ceiling effects may have impacted its effectiveness.

Several important conclusions are drawn from the study: first, as noted by the authors, use of one measure of fidelity may be insufficient in providing a complete picture. Though the SET is widely used and has been found in many studies described above to be related to behavioral and academic outcomes, scores were unrelated to outcomes in this study while the IPI, which provides different information organized around a stage theory of implementation, was significantly related to both behavior and academic performance (Pas & Bradshaw, 2012).

Additionally, though school contextual variables were controlled for in the final analysis, multiple factors were found to be related to both implementation fidelity and student outcomes. Years since training and percentage of certified teachers were significantly positively related to both IPI and BoQ scores. Student mobility was also significantly related to BoQ score. With regard to student outcomes, mobility level and school size were positively related to truancy and suspension rates respectively. Years since training, school size, and student mobility were also related to academic achievement. The findings of Pas and Bradshaw (2012) support the assertion of Domitrovich et al. (2008) that a variety of school factors can impact intervention implementation, including administrative leadership and support, school culture, climate, and characteristics. In the following section, literature will be reviewed regarding the important role that school context, specifically school setting and economic status, can play in student outcomes as well as quality of intervention implementation.

Role of School Context

School context may be thought of as the characteristics of an individual school, such as the setting and makeup of students and staff. Although a large number of these characteristics might be included in a discussion of school context, variables that have been included frequently in studies include the ethnic and socioeconomic makeup of a school, school location (urban, suburban, or rural), grades taught, and staff characteristics such as the degree of turnover (Stewart, 2003; Mendez, Knoff, & Ferron, 2002; Hope & Bierman, 1998, Heck & Mayor, 1993). School context has been found to be related to a number of academic and behavioral outcomes above and beyond student-level factors, and has also been tied to application of school discipline (Ma & Klinger, 2000; Bradshaw, Sawyer, & O'Brennan, 2009; Noltemeyer & McLoughlin, 2010). Given the evidence to support these relationships, an investigation of the possible moderating effects that school context may have on the effectiveness of PBIS is warranted. This review of literature is concentrated on the relationships between socioeconomic status (SES) and school location and academic, behavioral, and disciplinary outcomes.

Multiple studies have found relationships between school contextual variables and academic outcomes. School SES was found by Heck and Mayor (1993) to account for a large percentage of variance in academic outcomes in a study of elementary, middle, and high schools in a state in the western portion of the United States. Forced-entry multiple regression was used to identify effects of a number of school characteristics and academic indicators. The percentage of students receiving free or reduced lunch played an especially large role, accounting for 37% of variance in reading scores and 31% of variance in the area of math.

Barnes, Belsky, Broomfield, and Melhuish (2006) also studied school context, examining the extent to which neighborhood and school variables predicted academic achievement and school disorder. The authors studied schools in England that were identified using data from the Indices of Multiple Deprivation 2000 as having at least 10% of students who lived in areas of “social and economic deprivation”. In addition to neighborhood-level of poverty as identified by the Indices, variables of interest included school disorder (bullying, aggression between students, suspensions, parental aggression, and incidents reported to police), school achievement (percentage of 7 and 11 year olds meeting minimum standards on English, math and science achievement tests), and school characteristics (school size, percentage of students receiving free meals, percentage of students speaking a language in addition to English, and percentage of students receiving special education services). Stepwise linear regression indicated that the percentage of students receiving free lunch significantly predicted achievement scores in all academic areas for both 7 and 11 year olds above and beyond neighborhood conditions.

Other studies have also examined the effects of school context on academic outcomes, analyzing the extent to which context accounts for variance above and beyond individual variables. These studies indicate a complicated relationship between student-level variables, SES, and academic achievement. Caldas and Bankston (1997) conducted a four-step regression analysis to examine the extent to which overall academic achievement in Louisiana high schools is associated with peer poverty and social status when controlling for individual variables. When peer poverty status was added, it had a small but significant negative effect on achievement over and above individual variables. Peer family social status also had a small but significant effect when introduced in the third step, though the effect size

of peer poverty status was decreased when social status was included. In the final step of the analysis the percentage of black students was introduced, which had a significant negative effect on achievement. With the addition of ethnicity, peer poverty had a small positive effect, a change in direction the authors indicated may be due to the close relationship between peer poverty level and the percentage of black students.

Jargowsky and El Komi (2011) also found a positive relationship between the percentage of students receiving free and reduced price lunch and academic achievement when included in a regression model. Their study examined relationships between neighborhood, school, and student variables and average reading and math scores among 5th through 8th grade students in Texas. Neighborhood variables included poverty rate, percentage of children in married couple families, and percentage of adults with college degrees. School variables included the percentage of students receiving free or reduced price lunch and student turnover, and student variables included ethnicity and previous performance on achievement tests. Four value-added regression models were created, first with student variables, then with student and school variables, followed by student and neighborhood variables, and finally all three sets of variables. In both models that included school-level variables, there was a significant positive relationship between the percentage of students with free or reduced price lunch and reading and math scores. The authors noted, however, that the relationship changed when average peer score was removed from the model.

Ma and Klinger's (2000) study on the effects of individual and school factors on Canadian 6th grade students' academic achievement did not replicate the findings of Jargowsky and El Komi (2011) regarding the relationship between SES and academic outcomes.

Variables of interest were socioeconomic status as measured by students' responses when asked about educational items in the home and family activities, ethnicity, the number of parents in the home, and number of siblings. Additionally, parental involvement, school disciplinary climate, and academic press were included. Two-level hierarchical linear modeling was used to analyze effects, with student-level variables included in the first level and both student and school variables included in the second level. Multiple school variables had significant effects: school SES had a positive effect over and above student-level variables in the areas of math, reading and writing.

The previous studies indicate a complex interplay between school SES, other contextual variables, and academic outcomes. The strength and even direction of the relationship between school SES and academic outcomes varies depending on what other variables are included in the model. Sirin (2005) found that school setting may be another important variable to consider: a meta-analysis regarding SES and academic achievement found that school setting acted as a moderator of the effects of family SES. Effect size was greatest in suburban and lowest in rural schools. With regard to the relationship between context and perceptions of the school environment, behavior, and discipline, evidence is less ambiguous, and there is a greater breadth of literature examining the role of not only school SES but setting as well.

Anderman (2002) examined the effects of student- and school-level variables on school belongingness and other psychological outcomes among students who were tracked as part of the National Education Longitudinal Study. Student variables included gender, ethnicity, academic achievement, and socioeconomic status, while school variables included school setting (urban, suburban, or rural) and size. Hierarchical linear modeling was used to

analyze relationships: school location had significant effects in two of the three areas of school belongingness that were assessed. Students in urban schools were significantly more likely to report being victimized than students in suburban schools. Students in both urban and rural schools were significantly more likely than suburban students to perceive their school as unsafe.

Bradshaw, Sawyer, and O'Brennan (2009) also investigated school contextual variables, examining the extent to which indicators of school disorder were associated with bullying and attitudes about bullying among elementary and middle school children in Maryland. Indicators of disorder included student-teacher ratio, student poverty level (percentage of students receiving free or reduced-price lunch), suspension rate, student mobility, and school location (urban or suburban). Multilevel modeling was used to analyze the data. The model that included both student-and school-level variables produced no significant school-level variables related to victimization at the elementary level. With regard to endorsement of a retaliatory attitude, the percentage of students receiving free and reduced-price lunch was associated with increased odds of student endorsement of retaliatory attitudes. Additionally, free and reduced-price lunch was associated with decreased odds of students reporting feeling safe at school. Similar to Anderman's (2002) findings, Bradshaw et al. (2009) found that there was a significant association between schools' suburban location and increased odds of feeling safe among elementary students.

Stewart (2003) investigated the extent to which school culture, organizational structure, and social milieu were linked to misbehavior among high school students. Independent variables of interest included level of school attachment, commitment, involvement, belief in school rules, proportion of nonwhite students, school size, school

poverty level, and location. The dependent variable, misbehavior, was measured through student self-assessment of the number of physical altercations, violations of school rules, and suspensions that occurred during the first half of the school year. Contrary to their hypothesis that all of these variables would be linked to the level of misbehavior, only school size and location were significant predictors. Large schools and urban location were significantly related to higher levels of student misbehavior.

Hope and Bierman (1998) focused on school setting, examining differences in home and school behavior patterns among kindergarten students in rural and urban settings. They analyzed parent and teacher ratings of behavior using a repeated measures ANCOVA procedure, finding that rural and urban parents identified similar levels of behavior problems, but that teachers in urban settings identified more externalizing problems at school. When the analysis was confined to children with elevated behavior problems, location again was tied to significant differences: a chi-square test indicated that children in urban settings were more likely to exhibit problems only in school, while rural children were more likely to exhibit problems only at home. The authors' analyses also indicated that differences existed in the types of problems among children in each setting. Children in urban schools were more likely to exhibit externalizing or comorbid problems, while children in rural schools were equally likely to exhibit externalizing, internalizing, or comorbid problems.

While the previous studies focused on student behaviors and attitudes, other authors have examined school context as it relates to suspensions. Rate of suspensions reflects not only student behavior, but also staff attitudes and disciplinary approaches. As discussed earlier, the reliance on suspension in schools in the United States has been seen as ineffective and problematic by many. The American Psychological Association Zero Tolerance Task

Force (2008) noted that, rather than acting as a deterrent of future misbehavior, reception of suspension predicts higher rates of suspension in the future. Exclusionary discipline has also been applied disproportionately to black students, and used for subjective offenses such as disobedience (Forsyth, Howay, Pei, Forsyth, Asmus, & Stokes, 2013; Gregory, Skiba, & Noguera, 2010). These issues have served as an impetus for development of alternative approaches to behavior management: the statewide rising rates and disproportionate application of suspension were cited as a reason for North Carolina's development of a PBIS initiative (Irwin & Algozzine, 2008).

Research by Christie, Nelson, and Jolivette (2004) indicated that school context is related to the application of exclusionary discipline. They examined contextual variables related to rates of suspension in middle schools in Kentucky. Variables included the number of school board and law violations, school size, attendance rate, academic achievement, retention rate, percentage of males, SES as measured by the percentage of students receiving free or reduced-price lunch, percentage of white students, amount of spending per student, student-teacher ratio, and average teacher salary. Among other variables, they found that the percentage of students receiving free or reduced-price lunch was significantly positively correlated with suspension rate. The authors also performed a between-groups MANOVA to compare characteristics of the 20 schools with the lowest and 20 schools with the highest suspension rates. They found significant differences between the groups for 8 of the variables, including the percentage of students receiving free or reduced-price lunch.

Percentage of free and reduced price lunch was also associated with suspension rates in a study by Mendez, Knoff, and Ferron (2002). They examined out-of-school suspension patterns in a large, diverse school district in Florida. To identify related contextual variables,

the authors computed Pearson product moment correlations between suspension rates and a number of variables, including percentage of Black, Hispanic, and White students, academic achievement, class size, and the percentage of students receiving free and reduced-price lunch. Variables that were significantly positively associated with suspension rates at both the elementary and secondary level included the percentage of black students, mobility rate, and percentage of students receiving free or reduced-price lunch.

Even after controlling for risk factors, Hemphill et al. (2010) found an association between SES and suspension rates in middle schools in Australia. The authors assigned schools to one of four groups depending on which quartile of SES their neighborhood fell in. Logistic regression was used to control for the risk factors of poor family management, academic failure, and interaction with antisocial peers. Results indicated a significant association between area SES and suspension rates. Students in schools in the lowest quartile of SES were significantly more likely to receive a suspension than students in schools in higher quartiles. As the level of area SES increased, the suspension rates decreased: the authors found that students in schools in the second and third quartiles were one-third less likely to receive a suspension, and that students in the highest quartiles were two-thirds less likely.

Noltemeyer and Mcloughlin (2010) investigated the roles of school location and income level in disproportionate discipline practices. They examined the extent to which school setting (rural/agricultural, rural/small town, urban, major urban, urban/suburban), school income level (very high poverty, low, low to moderate, moderate to high, and very high median income) and student ethnicity (Black and White) were associated with rates of exclusionary discipline. A MANCOVA and univariate ANCOVAs were conducted to

investigate differences. The proportion of economically disadvantaged students in each district was included as a covariate in order to control for SES. MANCOVA results revealed a significant interaction between location and school income level and ethnicity. Post hoc analyses found the greatest disproportionality of application of exclusionary discipline in major urban schools with very high levels of poverty, and the lowest in rural/agricultural schools with low poverty.

Given the number of studies providing evidence to support the relationship between school context and student behavior and discipline, it is not surprising that evidence also suggests a relationship between school context and quality of implementation of prevention programs. Payne and Eckert (2010) examined the extent to which the characteristics of program providers and program structures, school climate, and school and community structure impacted the implementation of a number of different prevention programs. Specifically, the authors studied implementation intensity, measured by the programs' level of use, frequency of operation, duration, number of lessons or sessions, and frequency of student participation. School and community factors that were included in the study were community poverty, school size and urbanicity, grade level, and teacher turnover.

Using two-level hierarchical linear modeling, the authors found that urbanicity, poverty, and school size significantly impacted the frequency of operation of prevention programs. In contrast to what might be expected, larger schools in urban environments with higher levels of poverty were associated with higher quality implementation. This finding was consistent with previous work by Payne, Gottfredson, and Gottfredson (2006), which the authors hypothesized could be explained either by the low response rate among urban

schools or as the result of the availability of federal funding tied to implementation of evidence-based programs.

Several studies have found that school context is related to fidelity of PBIS implementation specifically. In addition to the work of Pas and Bradshaw (2012) cited earlier indicating a relationship between contextual variables and schoolwide PBIS fidelity, Pas, Waasdorp, and Bradshaw (2014) found evidence that school context is also associated with classroom implementation. Classroom implementation was measured using the EBS. A longitudinal three-level hierarchical linear modeling procedure was used to analyze effects: in the third level, school-level variables were included, including PBIS status, mobility rate, percentage of Black students, suspension rate, school-wide EBS score, and baseline SET score. Many school-level variables were significantly associated with classroom implementation, including student-to-teacher ratio, percentage of Black students, and baseline SET score, which had positive relationships, and suspension rate, which was negatively related. The authors noted that the strength of the relationships was moderate.

Taken as a whole, the research appears to support the implementation of PBIS as a means of promoting improved student outcomes. At the same time, research indicates that the behavioral and academic outcomes frequently targeted by those implementing PBIS are also tied to complex relationships between school context and fidelity of implementation. The state of North Carolina has supported the implementation of PBIS in its schools for a number of years, and has developed mechanisms to assess and promote implementation fidelity and ascertain outcomes associated with PBIS. The following section will detail the scaling-up process of the PBIS initiative in North Carolina and identify areas requiring additional investigation.

PBIS Implementation in North Carolina

When PBIS was first implemented in schools in North Carolina in 2000, considerable planning went into ensuring quality implementation and development of support systems. The lead agency in this initiative, the Behavioral Support Services Section in the Exceptional Children Division, began with a single Positive Behavior Support Center in Durham, North Carolina (Irwin & Algozzine, 2005, p. 4). The center was created using IDEA funding, and technical support was provided by partners of the National Technical Assistance Center on Positive Behavior Interventions and Supports (Irwin & Algozzine, 2005, p. 3). The initial goals of the center were to implement core elements of PBIS such as the use of screening tools to identify students in need of intervention, promotion of social skills through universal teaching, and the application of a tiered model of prevention and intervention (Irwin & Algozzine, 2005, p. 5). Additionally, the center sought to develop and implement core elements of the support system including in-service training for all staff.

To begin scaling up application of the PBIS model throughout the state, nine additional schools were designated Regional Positive Behavioral Support Programs in counties across North Carolina in 2002 (Irwin & Algozzine, 2005, p. 6). Additional schools volunteered in 16 counties the following year. As the number of schools implementing PBIS grew, multiple measures were taken to promote standardization of the support that schools were receiving. The state conducted “training of the trainers” sessions, in which individuals who were responsible for training and coaching within their districts were provided with centralized training (Irwin & Algozzine, 2005, p. 9). School teams also received training using centrally-developed materials that had been adapted from another implementing state. Training was broken up into two modules, the first of which focused on universal aspects of

PBIS, followed by small group and targeted interventions. In addition to tracking the number of schools participating in training, the state also collected data from specific schools to monitor the extent to which the schools were implementing primary, secondary, and tertiary interventions, analyzing data on a regular basis, and tracking outcomes such as ODR and suspension rates (Irwin & Algozzine, p. 11).

As might be expected, the interest in PBIS and resources available to support its implementation varied across the state in those initial years. While some districts had yet to begin training and implementation, several embraced PBIS and allocated resources above and beyond what was provided by the state. Durham County and Charlotte-Mecklenburg Schools were cited by Irwin and Algozzine (2005, p. 8) as two such districts. Durham hired a PBIS coordinator, reallocated staff to support implementation, and provided additional training to schools and individual teachers. Similarly, Charlotte-Mecklenburg staff received training from a partner of the National Technical Assistance Center on PBIS as well as weekly support provided by four full-time coaches. By the 2005-2006 school year, the number of full-time coaches grew to 11 (Irwin & Algozzine, 2007, p. 10).

Though PBIS implementation was not as widespread in most districts, 59 of the 115 districts in the state had PBIS schools by the 2004-2005 school year (Irwin & Algozzine, 2006, p. 8). The state allocated over \$300,000 to support PBIS, which helped to provide regional training opportunities for school teams and coaches. These opportunities consisted of six days of training spread out throughout the school year (Irwin & Algozzine, 2006, p. 11). The state also continued to expand their analysis of fidelity of implementation during the 2004-2005 school year. The SET was used to evaluate the original 16 implementation sites, which indicated that nine schools met 80/80 criteria (scores of at least 80% in the areas

of Behavioral Expectations taught and Total Implementation). Fifteen of the schools had scores of at least 80% in the area of Total Implementation (Irwin & Algozzine, 2006, p. 12).

The following year, 14 of 35 schools that were evaluated with the SET by PBIS coordinators met 80/80 criteria (Irwin & Algozzine, 2007, p. 14). By the 2005-2006 school year, 296 schools in over half of the districts in the state participated in training and implementation (Public Schools of North Carolina, 2013). In subsequent years, both the number of schools participating in the PBIS initiative and the number of schools assessed using the SET grew. By the 2011-2012 school year, 1154 schools had participated in training. Seven hundred fifty eight of those schools were actively implementing PBIS, representing a significant portion of the 2512 schools in the state (Public Schools of North Carolina, 2013). Among those schools, 604 had fidelity of implementation assessed during the 2011-2012 school year using the SET. Average scores for these schools was over 80% in the area of Total Implementation.

Evaluation of quality of implementation was not limited to use of the SET, however. As theorized by Domitrovich, et al. (2008) and supported by Pas and Bradshaw (2012), many factors impact the quality of implementation of an initiative. These factors include the effectiveness of the intervention itself as well as the degree to which support systems are in place. Beginning in the 2007-2008 school year, North Carolina instituted a system to recognize high-quality implementation that culled information from a number of sources about fidelity of implementation, the degree to which support structures were in place, and whether outcomes were affected. The state differentiated between three levels of quality; schools meeting the lowest level of recognition were named Green Ribbon schools, schools meeting the next level of criteria were named Model schools, and schools meeting the

highest criteria were named Exemplar schools. Levels of recognition were based upon criteria in three areas: systems, practices, and data.

Multiple systems-level requirements were uniform across the three levels of recognition, including active administration support, the existence of a PBIS team that meets at least eight times per year, and an identified in-school coach (Public Schools of North Carolina, n.d1.). Training requirements were tied to the three levels of recognition: a requirement of Green Ribbon recognition was completion of the universal module, Model Schools were required to complete the universal and small group intervention modules, and Exemplar schools were required to complete a training on targeted interventions in addition to the first two modules. These requirements ensured that schools had received standardized training, had access to on-going coaching, and that ongoing evaluation of PBIS implementation within schools was taking place on a regular basis.

The recognition system reinforced implementation with fidelity by requiring fidelity checks completed by school teams themselves as well as assessments completed by outside observers. Recognition required submission of SET or BoQ scores completed by either regional PBIS consultants or district-level coordinators or coaches. Green Ribbon recognition required SET Total Implementation scores of at least 80% or BoQ scores of 70%. Model recognition required SET Total Implementation scores of at least 90% or BoQ scores of at least 80%, and Exemplar schools were required to meet at least 95% of SET Total Implementation criteria or 90% of BoQ criteria (Public Schools of North Carolina, n.d.).

Additionally, school teams completed the Implementation Inventory Online (IIO), an assessment that provides information about the extent to which practices, systems, and mechanisms for data-based decision-making are in place for primary, secondary, and tertiary

levels of PBIS (Hughes & Algozzine, 2008). Schools with a total score on the IIO of under 70% are identified as a start-up. Schools that have met at least 80% of the school-wide total but less than 80% of the secondary total are identified as Level 1 implementers. Level 2 status is assigned to schools that have met at least 80% of school-wide and secondary criteria but less than 80% of tertiary criteria. Level 3 status denotes attainment of at least 80% of school-wide, secondary, and tertiary criteria. Finally, schools that have met 80% of school-wide, secondary, and tertiary criteria as well as 80% of subscale criteria are identified as Level 4 schools (Hughes & Algozzine, 2008). Green Ribbon recognition required an attainment of Level 1 on the IIO, while Model recognition required attainment of Level 2. Exemplar status required attainment of Level 3 or 4 on the IIO (Public Schools of North Carolina, n.d.).

A final component of the recognition system was systematic data collection. In addition to assessment and reporting of fidelity using the tools described above, schools earning recognition were required to enter referral data into a state database. Further, Exemplar schools were required to produce evidence of improvement in behavioral and achievement indicators, as well as a further school-identified area, such as staff retention, school climate, or special education referral rate (Public Schools of North Carolina, n.d.).

The number of schools that were recognized for quality implementation of PBIS grew rapidly in the years following introduction of the program. In 2007-2008, the first year of the program, a total of 68 schools earned recognition (Green Ribbon: 46 schools, Model: 8 schools, Exemplar: 14 schools) (Public Schools of North Carolina, 2013). By the 2013-2014 school year, 513 schools earned recognition (Green Ribbon: 171 schools, Model: 199 schools, Exemplar: 143 schools) (“PBIS Green Ribbon Schools 2013-2014”, n.d., “PBIS

Model Schools 2013-2014”, n.d., “PBIS Exemplar Schools 2013-2014”, n.d.). This number represented 21% of all schools in North Carolina. An evaluation of PBIS in North Carolina indicated that recognition of quality implementation was linked to positive outcomes, including lower rates of out-of-school suspension and higher graduation rates among recognition schools, though the significance of observed differences was not examined (Public Schools of North Carolina, 2013).

CHAPTER 3: RATIONALE FOR STUDY AND RESEARCH QUESTIONS

Development of this dissertation was influenced by the multi-level model of factors that can impact implementation quality as defined by Domitrovich, et al. (2008). The authors defined implementation quality as the degree to which an intervention is delivered as planned. They noted that the intervention itself and the systems that support its implementation are two related but independent components. While much research is dedicated to interventions themselves, support systems provide an integral role in large-scale implementation: “to reduce variability in the quality of program implementation” (Domitrovich, et al., 2008, p. 7). The intervention and its support system together make up two layers that directly impact quality of implementation, and, as an extension, the outcomes that are obtained.

Quality of implementation is also impacted by factors at the macro-, school-, and individual-level (Domitrovich, et al., 2008). Macro-level factors include legislation, financing, and human capital. At the school level, factors include the culture, structure, and characteristics of schools, including school size, setting, and economic status. Individual-level factors include implementer characteristics and attitudes about the intervention. Analysis of PBIS implementation in North Carolina will provide information about whether the efforts by the state to create systems of support and recognition of quality implementation were associated with improved outcomes. Additionally, information will be obtained about the extent to which improved outcomes were obtained at schools with varying characteristics.

Research on behavioral and academic outcomes associated with PBIS has typically supported its use. Studies have found reductions in the rates of ODRs and suspensions, resulting in fewer hours of instruction lost by both students and teachers (Lassen et al., 2006; Muscott et al., 2008). The evidence to support a relationship between PBIS implementation and improved academic outcomes is more mixed, but some studies have found an improvement in reading and math achievement in PBIS schools. Many of the studies that examine PBIS outcomes have limitations, however, including small sample sizes and failure to employ tests of significance (Luiselli et al., 2005; McIntosh et al., 2011, Muscott et al., 2008). The purpose of this study is to systematically examine the relationship between quality of PBIS implementation, school economic status, school setting, and academic and behavioral outcomes.

To do so, this study will make use of data collected by the North Carolina Department of Public Instruction. As detailed above, PBIS implementation is widespread in the state. There are state and regional systems of support in place, as well as systems for quality assurance. Though evaluation reports published by the state examined outcomes, most reports did not examine significance of differences between or within schools. This study will compare outcomes between schools who were recognized as Green Ribbon, Model, and Exemplar schools. Schools will be grouped according to their overall economic status (low or moderate/high). Additionally, the number of schools in one of four settings (urban, suburban, town, or rural) will be identified. The number of schools who have earned PBIS recognition is not large enough to allow a three-factor analysis of variance: when a model was constructed that included PBIS recognition level, school setting, and school economic status, multiple cells had a small number of schools. For that reason, school setting is not

being included in the research questions. However, possible interactions between setting, economic status, and PBIS level will be explored and discussed.

Within the context of PBIS implementation in North Carolina, the state's initiative identified PBIS as a means of promoting improved academic and behavioral outcomes (Irwin and Algozzine, 2008, p. 2). This study will examine whether improved academic achievement and decreased use of exclusionary punishment are associated with the varying levels of PBIS recognition created by the state. Examining possible differences between schools earning different levels of recognition will provide information about the extent to which level of program quality is associated with improved outcomes. Examining possible differences between outcomes in schools differing in economic status will provide information about the effectiveness of PBIS in schools as a function of school characteristics. The following questions will be investigated to address gaps in the literature and a need for evidence regarding PBIS implementation within North Carolina. Under each question is listed the associated hypothesis:

- 1) Are there significant differences between schools in suspension rates of students as a function of quality of PBIS implementation and school economic status?

H₁: There will be significant differences in suspension rates of students as a function of the interaction of school economic status and PBIS recognition status.

- 2) Are there significant differences between schools in the percentage of students scoring at or above grade level on the EOG Reading Assessment as a function of quality of PBIS implementation and school economic status?

H₂: There will be significant differences in the percentage of students scoring at or above grade level on the EOG Reading Assessment as a function of the interaction of school economic status and PBIS recognition status.

- 3) Are there significant differences between schools in the percentage of students scoring at or above grade level on the EOG Math Assessment as a function of quality of PBIS implementation and school economic status?

H₃: There will be significant differences in the percentage of students scoring at or above grade level on the EOG Math Assessment as a function of the interaction of school economic status and PBIS recognition status.

CHAPTER 4: METHODOLOGY

Selection of Schools

A total of 180 elementary schools were selected for inclusion in the analyses associated with questions 1, 2, and 3. Selection was based upon the level of state PBIS recognition during the 2013-2014 school year (“PBIS Green Ribbon Schools 2013-2014”, n.d., “PBIS Model Schools 2013-2014”, n.d., “PBIS Exemplar Schools 2013-2014”, n.d.). The included schools served either pre-k through fifth grade or kindergarten through fifth grade students. Middle and high schools were excluded from this study as a relatively small number of these schools attained PBIS recognition. Schools from 44 districts were included in the study. Sixty elementary schools that received Green Ribbon recognition were selected for the Level 1 group. Sixty elementary schools that received Model recognition were selected for the Level 2 group, and sixty that received recognition as Exemplar schools were selected for the Level 3 group. Each of these groups were evenly divided between 30 “Low Economic Status” schools and 30 “Moderate/High Economic Status” schools. For cells in which more than 30 available schools were available, a random number generator was used to select schools for inclusion. Because setting was identified as a contextual variable that may impact implementation and outcomes, the number of schools in city, suburban, town, and rural settings were identified.

Procedure

Research questions 1, 2, and 3 were investigated through employment of a 3x2 factorial design. For each of the 180 schools, data were collected regarding PBIS recognition

status, school economic status, school setting, suspension rates, and academic achievement scores.

PBIS Level

As noted in the section above, there were three PBIS recognition groups. Recognition status was determined based upon a number of criteria related to systems, use of data, and practices (Public Schools of North Carolina, n.d.1). Systems-level requirements that were uniform across levels of recognition included an administration that actively supported PBIS, the presence of a PBIS team that met at least eight times per year, and an identified in-school coach. Additionally, Green Ribbon (Level 1) schools were required to have completed Training Module 1, Model schools (Level 2) were required to have completed Training Modules 1 and 2, and Exemplar schools (Level 3) were required to complete all three training modules.

Requirements related to use of data for each level of recognition included entering office referral data into a statewide tracking system. Schools earning recognition as an Exemplar school were required to provide documentation of positive trends in behavioral and achievement indicators as well as improvement in another school-identified area of need. Finally, schools were differentiated based upon fidelity of implementation. Schools were required to provide evidence of completion of the Implementation Inventory Online (IIO) by the school-based team, and evaluation of fidelity using the SET or BoQ (Public Schools of North Carolina, n.d.1). Green Ribbon schools were required to obtain a total score of at least 80% on the SET or 70% on the BoQ and a Level 1 on the IIO. Model schools obtained at least 90% of total criteria on the SET or 80% on the BoQ and a Level 2 on the IIO, and

Exemplar schools obtained at least 95% on the SET or 90% on the BoQ and a Level 3 or 4 on the IIO.

Measures Used for Identification of PBIS Recognition Level

School-Wide Evaluation Tool. The School-Wide Evaluation Tool (SET) is an instrument designed to provide information about fidelity of implementation of PBIS (Horner et al., 2004). It consists of 28 items, which are scored either 0 (not implemented), 1 (partially implemented), or 2 (fully implemented). Observations, interviews, and reviews of data, reports, and other PBIS-related materials are used to assign scores. A total score is provided, as are seven subscale scores. These seven subscales are based upon seven critical features of PBIS: definition of school-wide behavioral expectations, teaching of expectations, provision of rewards for following expectations, consistent implementation of a continuum of consequences for problem behavior, monitoring of problem behavior patterns and use of data for ongoing decision-making, active administrator support and involvement, and provision of district support through development of policy, training opportunities, and data collection methods.

The SET has high internal consistency ($\alpha=.96$) (Horner et al., 2004). In repeated administrations two to three weeks apart, there was an average of 97.3% agreement. Additionally, in 17 administrations of the SET that were scored by two separate observers, average item-by-item interobserver agreement was 99%. The SET also has demonstrated sensitivity: statistically significant differences were found when pre- and post-implementation scores were calculated (average pre-implementation score: 47.9%; average post-implementation score: 83.6%). A significant correlation was found ($r=.75$) when total

scores on the SET were compared with total scores on the EBSSAS, an instrument completed by school-based teams to assess PBIS implementation.

Benchmarks of Quality. The Benchmarks of Quality (BoQ) is a 53-item rating scale completed by school teams to provide information about fidelity of PBIS implementation. It provides scores in 10 areas: PBS Team, Faculty Commitment, Effective Discipline Procedures, Data Entry, Expectations and Rules, Reward System, Lesson Plans, Implementation Plans, Crisis Plans, and Evaluation (Cohen, Kincaid, & Childs, 2007). Although it has gone through a modest revision, the initial version of the BoQ had a high degree of internal consistency ($\alpha=.96$), test-retest reliability ($r=.94$), and interrater reliability ($r=.87$). There is a moderate correlation between the BoQ and SET ($r=.51$), which the authors theorized could be partially explained by the BoQ's increased ability to discriminate among high-fidelity schools (Cohen, Kincaid, & Childs, 2007). The authors also noted that larger decreases in ODR rates were observed among schools that received higher scores on the BoQ.

Implementation Inventory Online. The Implementation Inventory Online (IIO) is an instrument that allows schools to identify strengths and weaknesses of their PBIS implementation (Hughes & Algozzine, 2008). It is completed by school-based teams, who assign scores of 0 (not at all), 1 (partial), or 2 (full) to reflect the degree of implementation in each area. A total score is provided, as well as scores in the areas of practices, systems, and data-based decision making for primary, secondary, and tertiary levels of PBIS. These scores are used to identify the overall level of implementation. Start-Up status is assigned to schools with a total score below 70%. Level 1 status is assigned to schools meeting at least 80% of the school-wide total but less than 80% of the secondary total. Schools that have met

at least 80% of school-wide and secondary criteria but less than 80% of tertiary criteria are given Level 2 status. Level 3 status is assigned to schools that have met at least 80% of school-wide, secondary, and tertiary criteria. Finally, Level 4 status is assigned to schools meeting 80% of school-wide, secondary, and tertiary criteria as well as 80% of the Practices, Systems, and Data subscale criteria. Data related to reliability and validity of the IIO are not available.

School Economic Status

With regard to school economic status, each school was placed into a “Low Economic Status” or “Moderate/High Economic Status” group according to criteria used by Sirard, Ainsworth, McIver, and Pate (2005). Schools in which 67% or more of students received free or reduced-price lunch during the 2013-2014 school year were placed into the Low Economic Status group, while schools with less than 67% of their students receiving free or reduced-price lunch were placed into the Moderate/High Economic Status group. The percentage of students receiving free and reduced-price lunch at each school during the 2013-2014 school year was obtained through information published by the North Carolina Department of Public Instruction (n.d.). Each PBIS group included 30 Low Economic Status schools and 30 Moderate/High Economic Status schools.

Setting

The setting of each school included in the study was identified using information provided by the National Center for Education Statistics (NCES). The NCES uses census information to place schools into categories (Rural, Town, Suburban, and City). For the purposes of this study, schools were assigned to one of four groups: Rural, Town, Suburban, and City, based upon their NCES setting description for the 2013-2014 school year. The

NCES defined the City setting as inside an urbanized area and a principal city (National Center for Education Statistics, n.d.). The Suburb setting was defined as being outside a principal city and inside an urbanized area, and the Town setting was defined as being inside an urban cluster and outside of an urbanized area. The Rural setting was defined as being outside of an urbanized area and urban cluster.

Outcome Measures

Suspension Rates. Suspension rates and test scores for each school in the study were obtained through the North Carolina Department of Public Instruction website, which provides outcome information for each school in the state each school year. Data that were collected pertained to the rate of out-of-school suspensions per 100 students during the 2013-2014 school year.

End-of-Grade Assessments. Academic achievement was evaluated using the percentage of students who achieved at least proficiency on the North Carolina End-of-Grade (EOG) English Language Arts/Reading and Mathematics tests. The EOGs are administered to students beginning in third grade to gain information about students' attainment of knowledge and competencies related to the grade level curriculum in each subject. The EOG English Language Arts/Reading assessment has demonstrated internal consistency, with Cronbach coefficient alpha on the three forms of the assessment ranging from .88 to .92 (North Carolina Department of Public Instruction, 2014). The EOG Mathematics assessment has a similar degree of internal consistency, with Cronbach coefficient alpha ranging from .91 to .92 (North Carolina Department of Public Instruction, 2014). Additionally, content validity was ensured through the test creation process which involved development of test items by North Carolina teachers with knowledge of the curriculum and a review process

completed by additional teachers, content experts, and item editors (NCDPI/North Carolina Testing Program, 2014, p. 6). The percentage of students who demonstrated proficiency during the 2013-2014 school year was identified for each of the 180 schools included in the study.

Analysis

Following assignment of schools to each group and collection of suspension rate and reading and math EOG data, statistical analyses were conducted to address research questions 1, 2, and 3. For each of these questions, a two-way ANOVA was conducted to determine the presence of significant main effects and interactions. Prior to the ANOVA, Levene's Test for Homogeneity and the Shapiro-Wilk test were applied to assess whether the assumptions of homogeneity of variance and normality were met. With regard to the suspension data, these assumptions were not met. Examination of the data and the associated schools was conducted to assess accuracy of data input and appropriateness of inclusion of outlier schools. The values of suspension rates associated with outlier schools were adjusted to attempt to minimize the extent to which they skewed the data. The assumptions of ANOVA were met for the reading and math achievement data. No significant interactions were found between PBIS level and the three dependent variables. Significant main effects of the independent variables were identified for each of the three dependent variables, and effect sizes were calculated. Additionally, Tukey's HSD procedure was implemented to identify significant differences between cell means where appropriate.

To explore the possible interaction between PBIS level, school economic status, and school setting on behavioral and academic outcomes, an exploratory analysis was conducted with 288 pre-k-5 or k-5 schools that earned PBIS recognition during the 2013-2014 school

year. For each of the dependent variables of suspension rate per 100 students, percentage of students achieving proficiency on the EOG Reading Assessment, and students achieving proficiency on the EOG Math Assessment, the mean, standard deviation, and number of schools was identified for schools grouped by PBIS level, school economic status, and setting. The size of many of the groups was small; for the purposes of conducting a three-way ANOVA to identify possible relationships that bear further investigation, the variable of setting was condensed into two groups: City/Suburb and Town/Rural.

For each of the variables of suspension rate and percentage of math and reading proficiency, Levene's Test and the Shapiro-Wilk test were conducted to assess homogeneity of variance and normality. Similar to the procedure used in analysis related to Research Question 1, outliers were adjusted within the suspension data in an attempt to minimize their skewing effect. A three-way ANOVA was conducted for each dependent variable. Where interactions were found, contrasts were calculated to identify simple main effects. To reduce the risk of a Type I error, the total error rate of .10 for the two main effects was divided between each of the five simple main effects that were tested, as described by Ware (2014, p. 298). Thus, the significance level for each simple main effect was .02. For interactions involving PBIS, which had three levels, Tukey's HSD was used to identify significant differences between levels. Additionally, the Tukey-Kramer method was used to account for the unequal group sizes. Tukey's HSD was also used to identify significant differences between levels of PBIS where no significant interaction was present. Again, the Tukey-Kramer method was applied due to the presence of unequal groups.

CHAPTER 5: RESULTS

From a pool of 288 k-5 schools who earned PBIS recognition in 2013-2014, 180 schools were selected for inclusion in the analyses of the three research questions. A random number generator was used to select schools to be placed in one of six groups depending on the level of PBIS recognition earned and the economic status of the school. Though school setting was not included as a variable in the statistical analysis pertaining to the research questions, it was a variable of interest in this study. The distribution of schools across settings in each group is presented in Table 1. In order to investigate the three research questions listed below, analyses were conducted using R statistical software (Version 3.3.2).

Table 1: Number of school settings by group

		PBIS Level 1	PBIS Level 2	PBIS Level 3	Economic Status Totals
Low Economic Status	City	13	7	7	27
	Suburban	3	5	4	12
	Town	2	6	4	12
	Rural	12	12	15	39
Moderate/High Economic Status	City	10	7	11	28
	Suburban	6	6	9	21
	Town	2	1	0	3
	Rural	12	16	10	38
PBIS Level Totals		60	60	60	180

Research Question 1

The first research question pertained to the effects of economic status and PBIS recognition level on short-term suspension rates at k-5 schools in North Carolina: Are there significant differences between schools in suspension rates of students as a function of quality of PBIS implementation and school economic status? As can be seen in Table 2, there were differences in suspension rates between the levels of PBIS recognition. There was a small difference between Level 1 and Level 2, with a larger difference between Level 2 and Level 3. Average suspension rates were much lower among schools in the Moderate/High Economic Status group (2.89 per 100 students) than they were in the Low Economic Status group (9.98 per 100 students). For comparison, the average short-term suspension rate for elementary schools during the 2013-2014 school year was 13.12 per 100 students. Much higher standard deviations were seen in the Low Economic Status groups, particularly in the PBIS Level 1 and Level 2 groups.

Table 2: Short-term suspension rates per 100 students by economic status and PBIS level

		PBIS Level 1	PBIS Level 2	PBIS Level 3	Economic Status Means
Low Economic Status	Mean	10.76	12.25	6.93	9.98
	Median	6.20	5.03	4.87	
	sd	10.74	15.67	6.70	
	n	30	30	30	
Moderate/High Economic Status	Mean	3.55	2.62	2.50	2.89
	Median	2.45	1.57	2.01	
	sd	3.49	3.08	2.59	
	n	30	30	30	
PBIS Level Means		7.16	7.43	4.72	

Tests of Homogeneity and Normality. A two-way ANOVA was selected as the method of analysis to test the hypothesis associated with Research Question 1. Levene's Test for Homogeneity was applied prior to conducting the analysis. Results indicated that the variances were significantly different, with an F value of 6.2644. The significance level was 2.272e-05, below the critical value of .05. The assumption of normality was also tested using the Shapiro-Wilk test. For each of the six groups, the p-value was below .05. The results of these tests indicated that the heterogeneity of variance and non-normality of the data were areas of concern.

A visual inspection of a box-and-whiskers plot of the six groups indicated that there were outliers in four of the groups. Data entry for each of these schools was double-checked, and the schools were researched to determine whether there were any aspects that should exclude them from the study. The data entry was accurate, and an investigation of suspension rates in years before and after the year of interest generally found that the higher rates were consistent with the schools' trends. All of the outlier schools were regular, prek-5 or k-5 schools, indicating that they were part of the group of schools of interest. It was determined that there was no justification for excluding the schools, and in fact the high rates of these schools made it important to study them. While the highest rate of suspension per 100 students among the Moderate/High Economic Status schools was 13.75, 22 of the schools in the Low Economic Status group had higher rates. Seven of those schools had suspension rates above 30 per 100 students. Due to the positive skewness of the suspension data, median values for each group were calculated (see Table 2).

While the outlier schools were included in the analysis, an effort was made to minimize the extent to which they skewed the data. Using a technique described by Fidell

and Tabachnick (2003, p. 124), outliers were assigned new values. The first outlier was assigned a value of the first whole number after the suspension rate of the next highest school. Every outlier after that was assigned a value of one whole number greater than the previous value. This was continued until all of the outlier schools were given new values. Once the outliers were adjusted, descriptive statistics were recalculated. As can be seen in Table 3, the changes in means and standard deviations were small. The largest change occurred in the Low Economic Status PBIS Level 2 group. Levene's Test and the Shapiro-Wilk test indicated that homogeneity of variance and normality were still problems with the adjusted data set. Given the robustness of ANOVA with regard to these assumptions, a two-way ANOVA was used to test the hypothesis.

Table 3: Short-term suspension rates per 100 students by economic status and PBIS level with adjusted outliers

		PBIS Level 1	PBIS Level 2	PBIS Level 3	Economic Status Means
Low Economic Status	Mean	9.26	11.54	6.93	9.24
	sd	7.75	13.78	6.70	
	n	30	30	30	
Moderate/High Economic Status	Mean	3.55	2.26	2.48	2.76
	sd	3.49	2.03	2.53	
	n	30	30	30	
PBIS Level Means		6.41	6.90	4.71	

Hypothesis 1. Hypothesis 1 stated: There will be significant differences in suspension rates of students as a function of the interaction of school economic status and PBIS recognition status. Results of the two-way ANOVA did not indicate the presence of a significant interaction ($p = 0.170$) (see Table 4). There was no significant main effect of

PBIS level, either ($p = 0.226$). A significant effect of economic status was found ($F_{1,174} = 35.714, p < 1.26e-08$). The effect size of economic status on suspension rate was calculated ($\eta^2 = 0.1651$), indicating that approximately 16% of the variance of suspension rates was accounted for by economic status.

Table 4: Analysis of Variance: Short-term suspension rates per 100 students by economic status and PBIS level with adjusted outliers

	Sum of Squares	Degrees of Freedom	Mean Square	F-ratio	p-value
PBIS Level	159	2	79.5	1.502	0.226
Economic Status	1890	1	1889.8	35.714	1.26e-08
PBIS Level: Economic Status	189	2	94.6	1.787	0.170
Error	9207	174	52.9		

Research Question 2

The second question was in regard to the effects of economic status and PBIS recognition level on reading achievement in k-5 schools in North Carolina: Are there significant differences between schools in the percentage of students scoring at or above grade level on the EOG reading assessment as a function of quality of PBIS implementation and school economic status? Reading achievement was quantified as the percentage of students at each school who achieved grade-level proficiency or above on the EOG Reading Assessment. Table 5 provides information on the mean percentage in each of the six PBIS/Economic Status groups, the standard deviation in each group, and the marginal means.

The average percentage of students who passed the EOG Reading Assessment increased from the lowest to the highest PBIS Recognition Status groups. The difference in average percentage was relatively small between PBIS Level 1 and Level 2; there was a larger difference between Level 2 and Level 3. A large difference (64.08 to 44.63) was observed in the overall Moderate/High Economic Status group in comparison to the Low

Economic Status group. For comparison purposes, the percentage of students who attained proficiency on the EOG Reading assessment for the 2013-2014 school year was 56.3%.

When means for each of the six groups in the study were compared, the average percentage of students who demonstrated proficiency in Low Economic Status schools was higher at Level 2 than Level 1, and was again higher at Level 3 than Level 2. Among Moderate/High Economic Status schools, the percentage was slightly lower at Level 2 than Level 1, but was much higher at Level 3. Standard deviations of the Low Economic Status groups were higher at PBIS Level 1 and Level 2.

Table 5: Percentage of students who passed the EOG Reading Assessment

		PBIS Level 1	PBIS Level 2	PBIS Level 3	Economic Status Means
Low Economic Status	Mean	41.49	43.00	49.40	44.63
	sd	11.21	12.26	9.53	
	n	30	30	30	
Moderate/High Economic Status	Mean	62.39	61.56	68.31	64.08
	sd	7.59	9.25	9.85	
	n	30	30	30	
PBIS Level Means		51.94	52.28	58.86	

Hypothesis 2. Hypothesis 2 stated: There will be significant differences in the percentage of students scoring at or above grade level on the EOG reading assessment as a function of the interaction of school economic status and PBIS recognition status. In order to test the hypothesis, a two-way ANOVA was selected as the method of analysis. Levene's Test for Homogeneity was used to determine whether the assumption of homogeneity was met. The test produced an F-value of 1.5875. The significance level was 0.166, above the critical value of .05. It was also necessary to assess the normality of the data. This was

accomplished through use of the Shapiro-Wilk test. The p-value for each of the six groups was above .05, which indicated that the assumption of normality was met. The results of Levene's Test for Homogeneity and the Shapiro-Wilk test indicated that use of an ANOVA was appropriate to test the hypothesis. Results of the two-way ANOVA did not indicate the presence of a significant interaction between PBIS recognition and economic status ($p = 0.78866$) (see Table 6). Hypothesis 2 was not supported.

Table 6: Analysis of Variance: Percentage of students who passed the EOG Reading Assessment

	Sum of Squares	Degrees of Freedom	Mean Square	F-ratio	p-value
PBIS Level	1822	2	911	9.003	0.00019
Economic Status	17033	1	17033	168.328	<2e-16
PBIS Level: Economic Status	48	2	24	0.238	0.78866
Error	17607	174	101		

Although a significant interaction effect was not found for PBIS Level and Economic Status, a significant main effect was found of PBIS recognition status on reading achievement ($F_{2,174} = 9.003$, $p = 0.00019$) (see Table 6). The effect size of PBIS recognition status was calculated ($\eta^2 = 0.0499$). This indicated that approximately 5% of variance of EOG Reading Assessment performance was accounted for by PBIS recognition status. Post hoc analysis to determine significant differences between groups was conducted using Tukey's HSD. Results indicated that there was a significant difference between PBIS Recognition Level 1 and Level 3 and between Level 2 and Level 3 (see Table 7). A significant main effect on EOG Reading Assessment performance was also found for economic status ($F_{1,174} = 168.328$, $p < 2e-16$). The effect size of economic status was very large ($\eta^2 = 0.4665$). Economic status accounted for close to 47% of variance in reading performance.

Table 7: Tukey's HSD: PBIS Level and percentage of students who passed the EOG Reading Assessment

Comparison	Difference	Adjusted p-value
Level 2 - Level 1	0.342	0.9811037
Level 3 - Level 1	6.913	0.0006653
Level 3 - Level 2	6.572	0.0012975

Research Question 3

The third question examined the effects of PBIS recognition status and economic status on math achievement at k-5 schools in North Carolina: Are there significant differences between schools in the percentage of students scoring at or above grade level on the EOG math assessment as a function of quality of PBIS implementation and school economic status? The percentage of students who achieved grade level proficiency on the state EOG Math Assessment was used to evaluate math achievement. The mean percentage and standard deviation for each group is presented in Table 8, along with the marginal means. An examination of the marginal means for PBIS Recognition Status indicates that the percentage of students who passed the EOG Math Assessment was larger at higher levels of PBIS recognition. As was seen with reading achievement, the difference between Level 1 and 2 was small. A larger difference was found between Level 2 and Level 3. Again, similar to reading achievement data, examination of the Economic Status marginal means indicated a large difference between the Low (45.62%) and Moderate/High (64.51%) Economic Status groups, with a higher average percentage of students who passed the EOG found in the Moderate/High Economic Status group. For comparison, 51% of students statewide attained proficiency on the EOG Math Assessment in the 2013-2014 school year.

Table 8: Percentage of students who passed the EOG math assessment

		PBIS Level 1	PBIS Level 2	PBIS Level 3	Economic Status Means
Low Economic Status	Mean	42.84	43.97	50.06	45.62
	sd	11.16	13.83	12.16	
	n	30	30	30	
Moderate/High Economic Status	Mean	62.55	63.08	67.90	64.51
	sd	9.59	10.86	11.17	
	n	30	30	30	
PBIS Level Means		52.70	53.53	58.98	

Hypothesis 3. Hypothesis 3 stated: There will be significant differences in the percentage of students scoring at or above grade level on the EOG math assessment as a function of the interaction of school economic status and PBIS recognition status. A two-way ANOVA was selected to test Hypothesis 3. Levene's Test for Homogeneity was conducted to assess the degree to which variance was homogenous. An F-value of 0.975 was produced, with a significance level of 0.4346. This value was above the critical value of .05, indicating that the requirement of homogeneity of variance was met. The Shapiro-Wilk test was applied to assess the normality of the data. The p-value for each group was above .05, indicating that the assumptions for use of ANOVA were met. The two-way ANOVA did not indicate a significant interaction between PBIS Recognition Status and Economic Status ($F_{2,174} = 0.102, p = 0.90300$) (see Table 9). Hypothesis 3 was not supported by results of the analysis.

Table 9: Analysis of Variance: Percentage of students who passed the EOG Math Assessment

	Sum of Squares	Degrees of Freedom	Mean Square	F-ratio	p-value
PBIS Level	1398	2	699	5.253	0.00609
Economic Status	16052	1	16052	120.662	<2e-16
PBIS Level: Economic Status	27	2	14	0.102	0.90300
Error	23147	174	133		

Although a significant interaction effect was not found, the analysis revealed a significant main effect of PBIS Recognition Status on math achievement ($F_{2,174} = 5.253$, $p = 0.00609$) (see Table 9). Effect size was calculated ($\eta^2 = 0.0344$), indicating that approximately 3% of the variance of the percentage of students who passed the EOG Math Assessment was accounted for by PBIS Recognition Status. Post hoc analysis using Tukey's HSD was conducted to determine significant differences between levels of PBIS Recognition Status (see Table 10). A significant difference was not found between Level 1 and 2 but significant differences were found between Level 3 and the other two levels. A significantly higher average percentage of students achieved grade-level proficiency on the EOG Math Assessment in Level 3 than in Level 1 or 2. A main effect was also found for economic status, with results of the ANOVA indicating a significant effect of economic status on math achievement ($F = 120.662$, $p < 2e-16$). The effect size of Economic Status was large ($\eta^2 = 0.3951$), indicating that approximately 40% of variance in math achievement was accounted for by Economic Status.

Table 10: Tukey's HSD: PBIS Recognition Status and percentage of students who passed the EOG Math Assessment

Comparison	Difference	Adjusted p-value
Level 2 - Level 1	0.828	0.9182935
Level 3 - Level 1	6.282	0.0091112
Level 3 - Level 2	5.453	0.0279771

Outcomes among Schools Grouped by PBIS Level, Economic Status, and Setting

In order to explore possible relationships between PBIS Level, Economic Status, and Setting, descriptive statistics were calculated for suspension rates and reading and math achievement. Additionally, a three-way ANOVA was conducted for each dependent variable. Information about the setting, suspension rates, and percentage of students who achieved proficiency was gathered for all 288 prek-5 and k-5 schools who earned one of the three levels of PBIS recognition for the 2013-2014 school year. This included 82 Level 1 schools, 113 Level 2 schools, and 93 Level 3 schools. With regard to setting, schools were assigned to one of the following four groups: City, Suburb, Town, or Rural. In total, there were 95 City schools, 50 Suburb schools, 29 Town schools, and 114 Rural schools. When grouped by PBIS level, economic status, and setting, the number of schools in each group ranged from 1 (PBIS Level 2, Moderate/High Economic Status, Town) to 34 (PBIS Level 2, Low Economic Status, Rural). The number of schools in each group are listed in Table 11.

Suspension Rates. Table 11 also shows mean suspension rates and standard deviations among all 24 groups. Large differences were observed between PBIS Levels among the Low ES City schools and Low ES Suburb schools. For these groups, suspension rates were lower in schools in the higher PBIS Level groups. Low ES Town and Rural schools had lower suspension rates in PBIS Level 1, but did not have lower rates as PBIS Level increased. In comparison to their Low ES counterparts, the Moderate/High ES groups had much lower suspension rates. There were smaller differences in suspension rate as PBIS Level increased, and in the case of Moderate/High Town and Rural schools, rates were slightly higher among PBIS Level 2 and 3 schools than they were at Level 1.

Table 11: Suspension Rate per 100 students among schools grouped by PBIS Level, Setting, and Economic Status

			Low Economic Status	Moderate/High Economic Status
PBIS Level 1	City	Mean	13.12	5.53
		Sd	14.56	3.39
		N	25	10
	Suburb	mean	12.83	3.15
		Sd	12.18	4.03
		N	7	6
	Town	mean	8.25	3.03
		Sd	12.18	3.74
		N	4	2
	Rural	mean	7.22	2.19
		Sd	8.77	2.90
		N	16	12
PBIS Level 2	City	mean	11.09	3.47
		sd	12.33	3.54
		n	25	7
	Suburb	mean	10.41	3.30
		sd	6.37	4.17
		n	7	8
	Town	mean	15.48	5.26
		sd	17.49	-
		n	11	1
	Rural	mean	7.13	3.34
		sd	9.28	4.47
		n	34	20
PBIS Level 3	City	mean	4.06	3.07
		sd	4.04	2.53
		n	8	20
	Suburb	mean	6.19	1.29
		sd	8.38	1.08
		n	5	17
	Town	mean	7.59	4.71
		sd	6.49	4.21
		n	8	3
	Rural	mean	7.80	4.41
		sd	6.61	4.10
		n	16	16

In order to explore whether there were statistically significant interactions present between PBIS Level, Economic Status, Setting, and Suspension Rate, a three-way ANOVA was conducted. Because of the low number of schools in some groups, the variable of school setting was modified. Instead of four groups, schools were placed into one of two groups based upon their proximity to urban areas (City/Suburban or Town/Rural). Group sizes ranged from 13 to 45. As was expected, Levene's Test indicated that variances were significantly different between groups ($p = 7.124\text{e-}05$). The Shapiro-Wilk test also indicated normality was an area of concern: the p-value was below .05 in 9 of the 12 groups. An inspection of a box-and-whiskers plot indicated that there were a total of 17 outliers in those 9 groups.

To minimize the skewing effect of those schools, the technique described by Fidell and Tabachnick (2003, p. 124) that was used earlier in the study when analyzing suspension data was used again. Levene's Test continued to produce a significant F-value after adjustment of the outliers ($p = 2.525\text{e-}09$), and the Shapiro-Wilk test indicated that normality was a concern in 8 of the groups. Additionally, there was a fairly strong correlation of 0.589 between group size and standard deviation following adjustment of outliers. Given the exploratory nature of this portion of the study, a three-way ANOVA was used to analyze suspension data despite these concerns.

Results of the three-way ANOVA are presented in Table 12. The ANOVA indicated that there was not a significant interaction between PBIS Level, Economic Status, and Setting. There was a significant interaction between PBIS Level and Setting ($F_{2, 276} = 6.819$, $p = 0.001$). The effect size of the interaction between PBIS Level and Setting was calculated

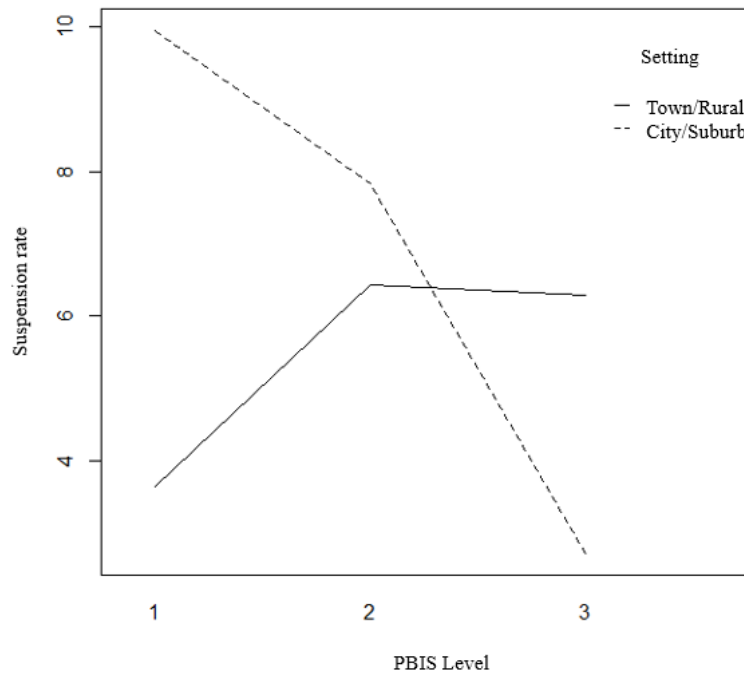
($\eta^2 = 0.039$), indicating that approximately 4% of variance of suspension rates was accounted for by this disordinal interaction.

Table 12: Analysis of Variance: Short-term suspension rates per 100 students by PBIS level, economic status, and setting with adjusted outliers

	Sum of Squares	Degrees of Freedom	Mean Square	F-ratio	p-value
PBIS Level	102.2	2	51.1	1.093	0.337
Economic Status	1377.4	1	1377.4	29.437	1.26e-07
Setting	87.5	1	87.5	1.871	0.172
PBIS Level: Economic Status	141.1	2	70.6	1.508	0.223
PBIS Level: Setting	638.1	2	363.1	6.819	0.001
Economic Status: Setting	43.3	1	43.3	0.926	0.337
PBIS Level: Economic Status: Setting	106.2	2	53.1	1.135	0.323
Error	12914.0	276	46.8		

Contrasts were calculated to identify simple main effects of PBIS Level and Setting on suspension rate. Setting was found to have a significant effect at PBIS Level 1 and Level 3 ($F_{\text{crit}} = 5.475$) (see Table 13). As can be seen in Graph 1, suspension rates were significantly higher in City/Suburb schools (mean = 9.96) than in Town/Rural schools (mean = 3.64) at Level 1. At Level 3, suspension rates were significantly lower in City/Suburb schools (mean = 2.71) than they were at Town/Rural schools (mean = 6.28). With regard to PBIS Level, contrasts indicated significant effects in the City/Suburb group ($F_{\text{crit}} = 3.968$). Post hoc analysis was conducted using Tukey's HSD, using the Tukey-Kramer method with a significance level of .02 to identify significant effects of PBIS Level. Significant differences were found between Level 3 (mean = 2.71) and the other two levels (Level 1 mean = 9.96; Level 2 mean = 7.83) in the City/Suburb group.

Graph 1: Interaction between PBIS Level and Setting on Suspension Rate



A significant main effect of Economic Status was also found ($F_{2, 276} = 29.437$, $p = 1.26e-07$). Effect size was calculated ($\eta^2 = 0.084$). This indicated that approximately 8% of the variance of suspension rate was accounted for by Economic Status.

Table 13: Simple main effects of PBIS Level and Setting on Suspension Rates

Simple main effect	F-ratio
Setting PBIS Level 1	16.99*
Setting PBIS Level 2	1.14
Setting PBIS Level 3	6.29*
PBIS Level City/Suburb	14.57*
PBIS Level Town/Rural	2.08

*Significant

Reading Proficiency. Table 14 shows the average percentage of students who demonstrated proficiency on the EOG Reading Assessment in schools grouped by PBIS Level, Setting, and Economic Status. Low ES Town schools had the lowest average proficiency rate in PBIS Level 1, but had the largest positive difference in percentage of

students demonstrating reading proficiency between PBIS Level 1 and Level 3 (+14.75). Low ES Suburban schools also had a fairly large positive difference in the percentage of students between PBIS Level 1 and Level 3 (+9.17). The other groups had smaller differences. In each setting, Moderate/High ES schools had a higher average percentage of students demonstrating proficiency than Low ES schools. However, among schools in the Town setting, the difference between Low ES and Moderate/High ES schools was much smaller at PBIS Level 3 (13.94) than it was at Level 1 (24.22). It should be noted that there were a small number of schools in each of the Town groups.

Table 14: Percentage of students who passed the EOG Reading Assessment among schools grouped by PBIS Level, Setting, and Economic Status

			Low Economic Status	Moderate/High Economic Status
PBIS Level 1	City	mean	37.82	63.40
		sd	8.72	9.08
		n	25	10
	Suburb	mean	44.23	66.32
		sd	7.73	9.17
		n	7	6
	Town	mean	35.08	59.30
		sd	8.40	3.82
		n	4	2
PBIS Level 2	City	mean	46.76	60.10
		sd	10.97	5.29
		n	16	12
	Suburb	mean	38.15	60.96
		sd	10.92	10.48
		n	25	7
	Town	mean	39.93	60.78
		sd	6.62	11.28
		n	7	8
PBIS Level 3	City	mean	45.99	53.60
		sd	17.32	-
		n	11	1
	Suburb	mean	48.13	61.07
		sd	9.90	7.83
		n	16	12
	Town	mean	35.08	59.30
		sd	8.40	3.82
		n	4	2
PBIS Level 4	City	mean	46.76	60.10
		sd	10.97	5.29
		n	16	12
	Suburb	mean	44.23	66.32
		sd	7.73	9.17
		n	7	6
	Town	mean	35.08	59.30
		sd	8.40	3.82
		n	4	2

		n	34	20
PBIS Level 3	City	mean	43.96	67.94
		sd	9.26	9.26
		n	8	20
	Suburb	mean	53.40	70.01
		sd	9.34	10.49
		n	5	17
	Town	mean	49.83	63.77
		sd	10.81	6.47
		n	8	3
	Rural	mean	51.50	64.41
		sd	8.72	11.41
		n	16	16

A three-way ANOVA was used in order to explore possible significant interactions between PBIS Level, Economic Status, and Setting. The variable of setting was condensed to two levels: City/Suburb and Town/Rural. Levene's Test for Homogeneity produced an F-value of 1.25 ($p = 0.254$), suggesting that the assumption of homogeneity of variance was met. The Shapiro-Wilk test also indicated that each of the 12 groups was sufficiently normal.

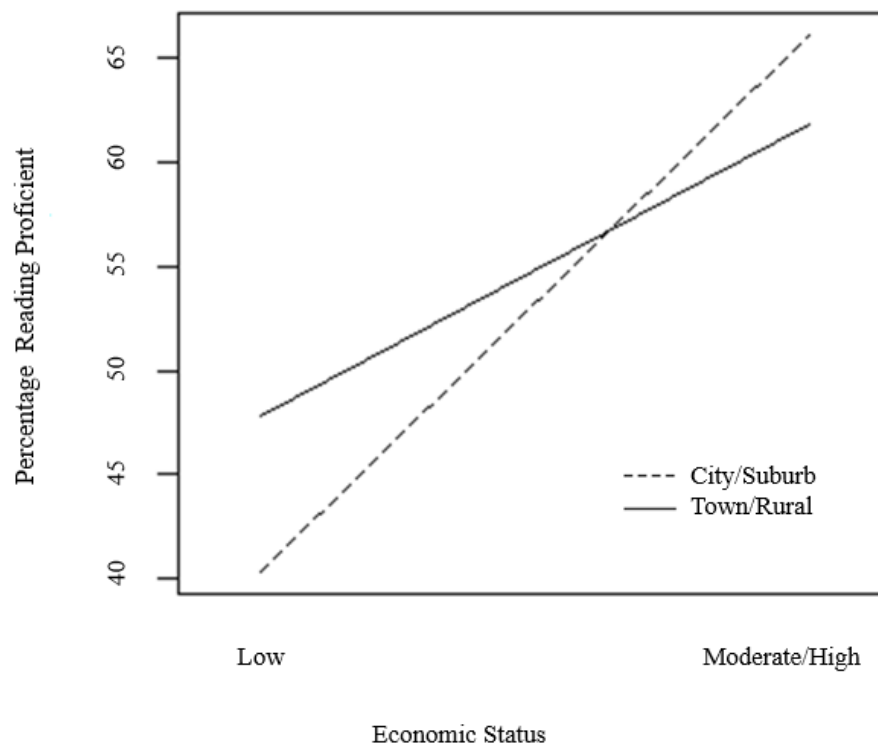
Table 15: Analysis of Variance: Percentage of students who passed the EOG Reading Assessment by PBIS Level, Setting, and Economic Status

	Sum of Squares	Degrees of Freedom	Mean Square	F-ratio	p-value
PBIS Level	1942	2	971	9.815	7.44e-05
Economic Status	21168	1	21168	214.574	<2.2e-16
Setting	121	1	121	1.228	0.269
PBIS Level: Economic Status	110	2	55	0.556	0.574
PBIS Level: Setting	324	2	171	1.641	0.196
Economic Status: Setting	1241	1	1241	12.578	0.0005
PBIS Level: Economic Status: Setting	8	2	4	0.043	0.958
Error	27227	276	98.6		

The ANOVA did not provide evidence of a significant interaction between PBIS Level, Economic Status, and Setting (see Table 15). There was, however, a significant interaction between Economic Status and Setting ($F_{2, 276} = 12.578$, $p = 0.0005$). The effect

size of the interaction was calculated ($\eta^2 = 0.02069$), indicating that approximately 2% of the variance was accounted for by the interaction. Contrasts were calculated to identify simple effects within the interaction (see Table 16). All conditions had F-values above the F_{crit} of 5.475, indicating that there were significant effects in all conditions. Graph 2 shows a disordinal interaction. At the Low Economic Status level, a higher average percentage of students achieved proficiency in Town/Rural schools than in City/Suburb schools. At the Moderate/High Economic Status level, a higher average percentage of students in the City/Suburb group achieved proficiency than in Town/Rural group.

Graph 2: Interaction between school setting and economic status on reading proficiency



A significant main effect of PBIS Level was also found. The effect size of PBIS Level was 0.032, suggesting that approximately 3% of the variance was accounted for by differences in PBIS Level. A post hoc analysis using Tukey's HSD indicated that there were

significant differences in the average percentage of students achieving reading proficiency between PBIS Level 3 (mean = 60.35) and the other two levels (PBIS Level 1 mean = 48.97; PBIS Level 2 mean = 49.24). The Tukey-Kramer method was used due to unequal group sizes.

Table 16: Simple main effects of Economic Status and Setting on EOG Reading Assessment Proficiency

Simple main effect	F-ratio
Setting Low Economic Status	23.72*
Setting Moderate/High Economic Status	5.623*
Economic Status City/Suburb	242.354*
Economic Status Town/Rural	66.742*

*Significant

Math Proficiency. The average percentage of students who achieved proficiency on the EOG Math assessment among schools grouped by PBIS level, setting, and economic status are shown in Table 17. The patterns observed are similar to those seen when reading proficiency was examined. Lower average levels of proficiency were seen in Low ES schools in each setting. The largest positive difference between PBIS Level 3 and PBIS Level 1 schools was again seen in Low ES schools in the Town setting (+9.58). Fairly large differences were also seen when Low ES Suburb schools were compared. Smaller positive differences were observed between PBIS Levels in the other settings.

Table 17: Percentage of students who passed the EOG Math Assessment among schools grouped by PBIS Level, Setting, and Economic Status

			Low Economic Status	Moderate/High Economic Status
PBIS Level 1	City	mean	39.70	63.38
		sd	10.17	10.28
		n	25	10
	Suburb	mean	44.43	65.35
		sd	7.58	13.48
		n	7	6

PBIS Level 2	Town	mean	37.03	60.20
		sd	11.84	0.99
		n	4	2
	Rural	mean	48.09	60.85
		sd	11.85	8.02
		n	16	12
	City	mean	41.09	60.40
		sd	11.82	12.17
		n	25	7
	Suburb	mean	43.41	65.91
		sd	7.09	9.45
		n	7	8
	Town	mean	45.93	63.70
		sd	18.47	-
		n	11	1
	Rural	mean	47.85	61.81
		sd	12.16	10.52
		n	34	20
PBIS Level 3	City	mean	44.33	67.61
		sd	8.53	9.87
		n	8	20
	Suburb	mean	52.56	69.41
		sd	12.79	11.56
		n	5	17
	Town	mean	46.61	65.70
		sd	13.00	8.91
		n	8	3
	Rural	mean	54.11	64.08
		sd	11.11	14.53
		n	16	16

Possible interactions between PBIS Level, Economic Status, and Setting on school math achievement were explored using a three-way ANOVA. Due to the small number of schools in some settings, schools were placed either into a City/Suburb or Town/Rural group. Levene's Test for Homogeneity produced an F-value of 1.31 ($p = 0.22$), and the Shapiro-Wilk test produced p-values above 0.05 for each group. The results of these tests indicated that the assumptions of homogeneity of variance and normality were met.

Table 18: Percentage of students who passed the EOG Math Assessment among schools grouped by PBIS Level, Setting, and Economic Status

	Sum of Squares	Degrees of Freedom	Mean Square	F-ratio	p-value
PBIS Level	1234	2	617	4.698	0.010
Economic Status	20139	1	20139	153.353	<2.2e-16
Setting	65	1	65	0.493	0.483
PBIS Level: Economic Status	52	2	26	0.197	0.820
PBIS Level: Setting	52	2	26	0.198	0.820
Economic Status: Setting	987	1	987	7.519	0.007
PBIS Level: Economic Status: Setting	5	2	2.5	0.019	0.981
Error	36245	276	131.3		

There was not a significant interaction between PBIS Level, Economic Status, and Setting on math achievement (see Table 18). A significant interaction was found between Economic Status and Setting ($F_{1,276} = 7.519$, $p = 0.007$) (see Graph 3). Effect size of the interaction was small ($\eta^2 = 0.0151$). Less than 2% of the variance of math achievement was accounted for by the interaction between Economic Status and Setting. To identify simple main effects within the interaction, contrasts were calculated (see Table 19). F_{crit} was 5.475. There were significant effects of Setting in the Low Economic Status group. Additionally, there were significant effects of Economic Status in both the City/Suburb and Town Rural groups.

There was also a significant main effect of PBIS Level ($F_{2,276} = 4.698$, $p = 0.010$). The effect size was 0.0189, indicating that less than 2% of the variance was accounted for by PBIS Level. Post hoc analysis was conducted using Tukey's HSD. There were significant differences between PBIS Level 3 (mean = 60.33) and the other two levels (Level 1 mean = 49.97; Level 2 mean = 50.56).

Graph 3: Interaction between Setting and Economic Status on math proficiency

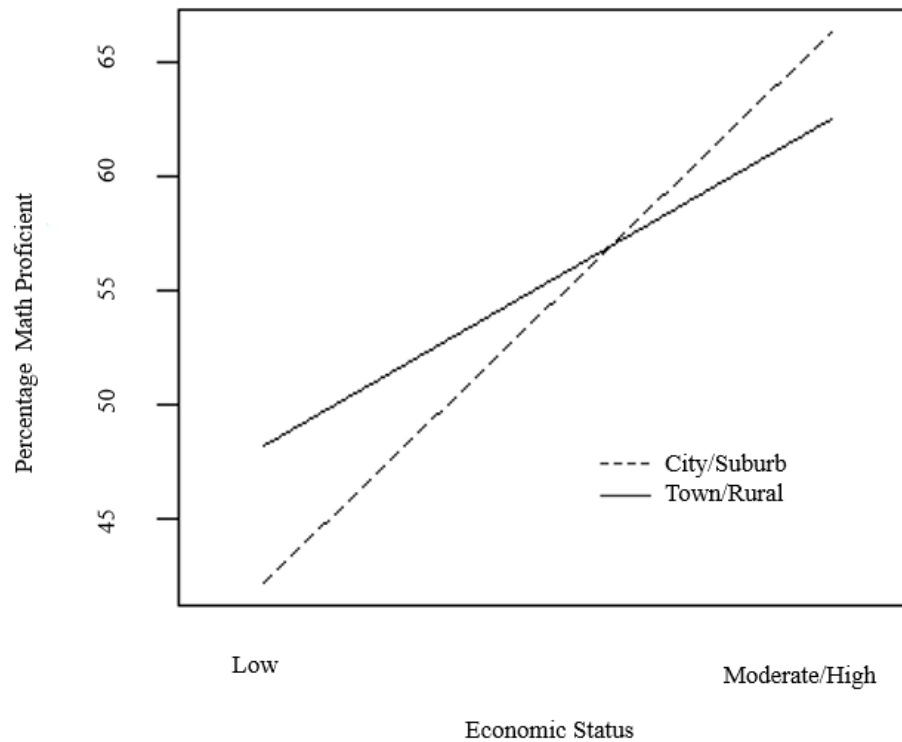


Table 19: Simple main effects of Economic Status and Setting on EOG Math Assessment Proficiency

Simple main effect	F-ratio
Setting Low Economic Status	11.127*
Setting Moderate/High Economic Status	3.377
Economic Status City/Suburb	159.203*
Economic Status Town/Rural	52.146*

*Significant

CHAPTER 6: DISCUSSION

Overview

The purpose of this study was to investigate the relationship between quality of PBIS implementation, school contextual variables, and behavioral and academic outcomes. PBIS is a framework that is used across the country to address schools' behavioral needs. The framework supports prosocial behavior and a safe school environment that promotes learning through employment of a three-tiered prevention model, application of evidence-based practices, and creation of systems to support and sustain implementation (Sugai & Horner, 2006). Given the link between behavior and academic performance, schools nationwide, including in North Carolina, have implemented PBIS as a means of improving not only behavioral outcomes but academic outcomes as well.

Many studies have found evidence to support a link between PBIS implementation and improved behavioral and academic outcomes (Lassen, et al., 2006; Muscott, et al., 2008; Luiselli, et al., 2005; Horner, et al., 2009). Research has also supported the theory posited by Domitrovich, et al. (2008) that outcomes of an intervention are tied not only to the soundness of that intervention but to multiple other factors, including fidelity of implementation and school-level factors (McIntosh, et al., 2011; Simonsen et al., 2012; Pas et al., 2014; and Pas and Bradshaw, 2012). To encourage high-quality implementation of PBIS, a system of recognition was created in North Carolina. Its method of identifying quality of implementation was more robust than the method used in many studies of PBIS. In addition to the use of two fidelity assessment instruments, the North Carolina system included level of

training, the presence of critical elements, and trends related to behavioral and academic outcomes.

In order to investigate whether different outcomes were obtained by schools attaining three different levels of implementation recognition, schools were placed in one of three groups based upon the level of PBIS recognition they earned from the state for implementation during the 2013-2014 school year. Schools were also defined as a Low Economic Status school or a Moderate/High Economic Status school based upon the percentage of students who received free and reduced-price lunch. Though not included in the inferential statistical analyses to investigate the three research questions, the setting of schools that earned recognition was also identified. Schools were assigned to one of four groups: City, Suburb, Town, or Rural. The behavioral outcome of interest in this study was the rate of suspension per 100 students of each school. This information was accessed through the North Carolina Department of Public Instruction website. The academic outcomes of interest were the percentages of students who achieved proficiency on the EOG Reading and Math Assessments. Again, this information was obtained through the Department of Public Instruction's website. The findings for each research question are detailed below.

Study Findings

Suspension Rates. The first research question posed was whether there were significant differences between schools in suspension rates as a function of quality of PBIS implementation and school economic status. The data related to suspension rates were non-normal, and there was greater variance in the Low Economic Status groups than in the Moderate/High Economic Status groups. Outlier suspension rates were adjusted to minimize

the extent to which they skewed the data, though problems remained even after this adjustment. An examination of median suspension rates of each group indicated the extent to which the rates of the Low Economic Status groups were positively skewed. While the median values of the Moderate/High ES groups were not drastically different from the means, the median values of the Low ES groups were lower than the means (the median suspension rate for PBIS Level 2 schools in the Low Economic Status group was 6.51 lower than the mean).

Results of a two-way ANOVA indicated that there was not a significant interaction between quality of PBIS implementation and school economic status on suspension rates. In addition, there was not a significant main effect of PBIS Level on suspension rates. This is in contrast to the findings of Simonsen et al. (2012) and Bradshaw, Mitchell, and Leaf (2010), who found a relationship between high-fidelity PBIS implementation and behavioral outcomes. Of note, Pas and Bradshaw (2012) did not find a relationship between fidelity as measured by the SET or BoQ and behavioral outcomes. The SET and BoQ were measures used in identification of levels of PBIS quality in this study. However, the recognition system used in this study was more robust, having other criteria, including another assessment of fidelity completed by school-based teams, level of training completed, and whether or not schools had demonstrated improvement in outcome data. Several observed patterns that may help to explain the lack of a relationship between PBIS Level and suspension rates may be related to the relatively low suspension rates observed at all levels of quality of PBIS implementation in the Moderate/High Economic Status group and high variability in the Low Economic Status group.

While the fact that there was not a significant relationship between PBIS Level and suspension rate was somewhat surprising, the results of the ANOVA with respect to Economic Status and suspension rate was less surprising. Consistent with the findings of Christie, et al. (2004), Hemphill, et al. (2010), and Mendez, et al. (2002), there was a large main effect of Economic Status on suspension rates. Low Economic Status schools had significantly higher suspension rates. Approximately 16% of the variance of suspension rates was accounted for by Economic Status. Of note as well were the large differences in variance between the Low and Moderate/High Economic Status groups. While almost all of the Moderate/High Economic Status school suspension rates were clustered below 10, the rates in the Low Economic Status schools ranged from 0 to 57, with almost 30 schools with rates above 10. These numbers indicate that, despite their attainment of PBIS recognition, suspension rates continued to be an area of concern for many of these schools.

Reading Proficiency. The second question posed was whether there was a significant difference between schools in the percentage of students scoring at or above grade level on the EOG Reading Assessment as a function of quality of PBIS implementation and school economic status. A two-way ANOVA was used to analyze this question. Results of the ANOVA indicated that there was not a significant interaction between quality of PBIS implementation and school economic status on reading achievement. The hypothesis that an interaction would be present was based upon previous studies regarding the effects of contextual variables on quality of implementation and the relationship between school socioeconomic status and behavioral and academic outcomes. Pas, et al. (2014) found that numerous contextual variables were related to classroom implementation of PBIS. Payne and Eckert (2010) investigated a variety of prevention programs, finding that urbanicity,

poverty, and school size were related to quality of implementation. Contrary to what might be expected, large, high-poverty schools in urban environments implemented at a higher level of quality. Schools with higher suspension rates and lower academic achievement have more room for improvement, so theoretically higher-quality PBIS implementation might be expected to pay greater dividends at these schools. However, this study did not find evidence of that interaction.

There were, nevertheless, significant main effects of PBIS level and economic status on reading achievement. School economic status accounted for a larger percentage of variance of average percentage of proficiency on the EOG Reading Assessment than PBIS Level. The significant effect of economic status is consistent with the findings of previous studies (Heck & Mayor, 1993; Barnes et al., 2006). Heck and Mayor's (1993) finding that 37% of variance in reading scores was accounted for by school socioeconomic status was similar to the finding of this study that 47% of variance was accounted for by economic status.

The amount of variance accounted for by PBIS level was 5%. Because of the nature of the study, causality cannot be inferred. However, differences in PBIS quality were associated with a small but significant positive difference in average reading proficiency. This supports the findings of previous studies that increased fidelity of PBIS implementation is associated with increased reading achievement (McIntosh, et al., 2011; Pas & Bradshaw, 2012). With respect to the small effect size, it is possible that, while there is a relationship between PBIS and academic achievement, the relationship may be relatively weak. PBIS is theorized to impact academic performance by increasing the instructional time of teachers, the academic engagement of students, and the orderliness and safety of schools. While

evidence supports this theory, there are many more factors that impact academic achievement, including quality of teachers, access to materials, and student-level factors. Perhaps quality behavioral practices allow for increased access to learning opportunities but are only one part of a complex puzzle.

Post hoc analysis indicated that there were significant differences in reading proficiency between Level 3 and the other 2 levels. While there was a very small difference between the average percentage in Level 1 (51.91) and Level 2 (52.22), a much larger difference was found between Level 2 and Level 3 (58.98). While causality cannot be inferred due to the associational nature of the study, a discussion is warranted of the differences in criteria between Level 3 and the other two levels that may possibly impact outcomes. All schools in the study had elements identified as critical to the success of PBIS, including an administration that was supportive of its implementation, a team that met at least eight times per year, and an in-school coach. There were differences in the training requirements, with primary-level training required of Level 1 schools, primary- and secondary-level training required of Level 2 schools, and primary-, secondary-, and tertiary-level training required of Level 3 schools. Each level of recognition was also associated with increasingly stringent fidelity requirements, with a requirement of 80% of total criteria on the SET at Level 1, 90% at Level 2, and 95% at Level 3. Finally, Level 3 required documentation of positive trends in behavioral and academic outcomes as well as a third area of the school's choosing.

It is impossible to isolate each of these elements to determine which may have an impact on academic outcomes. However, the differences in fidelity requirements were greater between Levels 1 and 2 than they were between Levels 2 and 3. Additionally, the

secondary-level training required of Level 2 schools would theoretically be expected to impact a larger percentage of students than tertiary-level training. However, as noted, the difference between reading outcomes of Level 1 and Level 2 schools was not significant, while the differences between those two levels and Level 3 were significant. It is possible that documentation of positive trends may have been a more important factor than differences in fidelity. It seems unlikely that a 5% increase in fidelity accounted for the significant difference in reading achievement observed between Level 2 and Level 3, especially given the lack of a significant difference between Level 1 and Level 2, which had a difference in fidelity requirements of 10%.

Fidelity of implementation may be necessary but not sufficient: it is possible that real differences in outcomes are not obtained until schools are actively engaged in data-based decision making. Of course, as noted above, this is speculation, as causality cannot be inferred from the data. There are alternative explanations for the differences in reading achievement between Level 3 and the other two levels. For instance, those schools may have had more effective administration or structures in place that allowed them to achieve both a higher level of fidelity and higher academic achievement. Further research is needed to investigate the extent to which higher fidelity and a focus on data are causally linked to higher levels of reading achievement outcomes.

Math Proficiency. The third question posed was whether there was a significant difference between schools in the percentage of students scoring at or above grade level on the EOG Math Assessment as a function of quality of PBIS implementation and school economic status. This question was evaluated using a two-way ANOVA. The results did not indicate the presence of a significant interaction between level of PBIS implementation and

school economic status. As noted in the discussion of reading proficiency, school contextual variables have been found to be related both to quality of implementation of prevention programs and to academic achievement (Payne & Eckert, 2010; Pas, et al., 2014). Though additional research in this area is needed, including the inclusion of additional variables such as school setting in the analysis, this study did not find evidence of an interaction between quality of PBIS implementation, economic status, and math achievement.

Significant main effects of PBIS implementation and school economic status were found for math proficiency. A much larger percentage of variance of average proficiency on the EOG Math Assessment was accounted for by school economic status. Similar to the findings on reading achievement, the findings of this study with regard to effects of school economic status on math proficiency were consistent with previous studies by Heck and Mayor (1993) and Barnes et al. (2006). Heck and Mayor (1993) found that school socioeconomic status accounted for 31% of variance, while school economic status accounted for 39% of variance in average percentage of students who passed the EOG Math Assessment.

A significant but limited amount of variance of the independent variable was accounted for by PBIS level. This outcome was consistent with previous findings by McIntosh, et al. (2011), Simonsen, et al. (2012), and Pas and Bradshaw (2012), which found a relationship between fidelity of implementation and math achievement. Of note, the fidelity requirements studied by McIntosh, et al. (2011) and Simonsen, et al. (2012) were lower than the levels that were examined in the present study. McIntosh, et al. (2011) considered schools that met between 70 and 79% of total criteria on the SET to be implementing with a moderate degree of fidelity, and schools that met 80% or greater of criteria to be implementing with high

fidelity. Simonsen, et al. (2012) used the 80/80 criteria, which considered schools that met 80% of criteria in the area of teaching expectations and 80% of total criteria on the SET to be implementing with high fidelity.

The present study investigated possible differences between schools implementing with a higher degree of fidelity, with the lowest level meeting 80% of criteria. The other two levels comprised schools that obtained increasingly higher degrees of high-fidelity implementation (90% and 95%). The study found differences in outcomes between the schools that obtained 95% of implementation and those that obtained 80% or 90%. As discussed in the reading proficiency section and noted below, however, the levels of recognition consisted of a more robust set of requirements that extended beyond a single measure of fidelity, limiting the ability to link the differences in outcome to differences in fidelity alone.

Post hoc analysis indicated that there were significant differences in the percentage of students attaining math proficiency between Level 3 and the other two levels. The differences between groups were similar to those found in the area of reading. While there was a small, statistically insignificant difference between Level 1 and Level 2 (52.92% and 53.39%), the average percentage of students who achieved math proficiency in Level 3 schools was a bit higher (59.06%). As noted in the discussion of reading proficiency, there were several differences between Level 3 and the other two levels, including the fidelity of implementation of PBIS and demonstration of positive behavioral and academic trends in Level 3 schools. The nature of the study does not allow for identification of a causal relationship: further research is needed to determine the extent to which fidelity of implementation and study of trends in data is causally linked to improved math outcomes.

School Setting. This study also investigated possible relationships between the independent variables of PBIS Level, Economic Status, and Setting and the dependent variables of suspension rate, reading proficiency, and math proficiency. This was accomplished through a series of three-way ANOVAs. Due to the relatively small number of schools in some of the groups, the schools which were originally placed into four distinct groups (City, Suburb, Town, and Rural) were combined into two groups (City/Suburb and Town/Rural). While no definitive conclusions can be drawn from the results of the ANOVAs, they indicate that further research in this area is needed.

With regard to suspension rate, an examination of the data prior to analysis indicated that variance was not homogenous across groups. Similar to the findings related to Research Question 1, there was higher variance in suspension rates in the Low Economic Status schools than the Moderate/High Economic Status schools. The data were also not normally distributed: most groups were positively skewed, and many had significant outliers. The values of these outliers were reduced to make them less extreme. A notable finding of the three-way ANOVA was the presence of an interaction between PBIS Level and Setting. While City/Suburb schools had a significantly higher mean suspension rate than Town/Rural schools at PBIS Level 1, that rate was significantly lower at PBIS Level 3. There was a significant difference of suspension rates between Level 3 and the other two levels in City/Suburb schools. It should be noted that while this interaction was significant, it accounted for only 4% of variance of suspension rate.

Previous research had found that there was a relationship between school setting and suspension rate: Noltemeyer and McLoughlin found that the highest rates of exclusionary discipline were found in large, high-poverty schools in urban settings. Research has also

found a relationship between context and PBIS implementation. Pas and Bradshaw (2012) found that context was related to schoolwide fidelity, and Pas, Waasdorp, and Bradshaw (2014) found a relationship between context and classroom implementation of PBIS. The findings of this study point to a continued need for research in this area. Specifically, the findings appear to indicate that higher quality implementation may be associated with significant positive differences with regard to suspension in schools in city or suburban settings, and with no significant differences in schools in less urban settings.

Regarding reading and math outcomes, a significant interaction between PBIS Level and Setting was not found. Instead, similar to the findings of this study regarding Research Question 2 and 3, significant main effects of PBIS Level were found on the percentages of students attaining proficiency on EOG Reading and Math Assessments. Again, the effects of PBIS on reading and math achievement were limited, accounting for 3% and 2% of variance of the two variables respectively. With regard to both math and reading, significant differences were found between PBIS Level 3 and Level 2. The findings were in line with previous studies that found that increased fidelity of PBIS was related to higher levels of reading and math achievement, and are consistent with the earlier findings of this study when only PBIS Level and Economic Status were included in the analysis (McIntosh, et al., 2011; Pas & Bradshaw, 2012; Simonsen, et al., 2012).

While not posed as research questions for this study, significant interactions were found between Economic Status and Setting on both reading and math achievement. The interactions accounted for approximately 2% of variance of reading and math proficiency. A disordinal interaction was found with respect to reading: a significantly higher percentage of students passed the EOG Reading Assessment in Low Economic Status Town/Rural schools,

while a higher percentage passed the assessment in Moderate/High Economic Status City/Suburb schools. A similar pattern was found with the math data, though the difference between settings was not significant at the Moderate/High Economic Status level.

The findings appear to suggest complex interplay between contextual variables. Further research that includes larger numbers of school in each setting and allows for comparison of schools in all four settings rather than the combined setting groups that were used in this study will help to clarify the relationships between context and achievement. What appears quite clear from previous studies and from this study is the relationship between school economic status and achievement (Heck & Mayor, 1993; Barnes et al., 2006). Examination of the interactions between economic status and setting found significant differences between Low and Moderate/High Economic Status schools in both setting groups with respect to reading and math proficiency.

Limitations

There are several limitations to this study that impact the extent to which conclusions can be drawn. The study made use of several sources of information, all of which were previously collected and publically available. This approach had several advantages, including the ability to assess the outcomes associated with relatively large-scale implementation of PBIS in a real-world setting as well as the effectiveness of a program designed to promote quality implementation.

At the same time, this use of existing data had several disadvantages. Schools were not randomly selected or placed into groups. The schools who chose to pursue training in, and implementation of, PBIS may or may not constitute a representative sample of elementary schools in North Carolina. It appears logical that schools that have behavior

problems may be more likely to pursue a positive behavior framework, but at the same time, it is also a possibility that schools that attained recognition may have a more driven and organized administration and staff. While the analyses indicated significant differences in academic achievement between the highest level of quality and the other two levels, the extent to which those differences were attributable to differences in PBIS quality is not known. There may have been preexisting differences in academic achievement between schools at the various levels prior to implementation of PBIS, or those schools could have been implementing other programs that explain the differences in outcomes.

Additionally, the use of the list of schools earning recognition had several limitations. Attainment of the various levels of recognition was not limited solely to fidelity: other factors, including the amount of training attained and the production of data indicating positive trends in outcomes, were also considered. This was positive in the sense that PBIS quality was based on multiple indicators. At the same time, the differences between the levels in these multiple aspects impact the ability to draw conclusions about the effect of any one of them. Additional research is needed to identify and define the relationships between these factors and outcomes of interest. The study also included fewer groups than originally intended. A group of schools that had not been trained in PBIS as well as a group of schools that had been trained but had not earned recognition were intended to be included. Inclusion of these groups would have provided a more complete picture of the role of PBIS implementation as a whole, and more specifically of the role of quality of implementation. Unfortunately, this information did not exist, as the list compiled by NCDPI of schools which had completed training was overwritten each year.

There were also several limitations with respect to the study of possible interactions between contextual variables and quality of PBIS. While research questions that incorporated both economic status and setting would have provided a more complete picture of the effects of school context, there were not enough schools to create groups divided by PBIS level, economic status, and setting with sufficient numbers of schools in each group. The exploratory analysis, which included two setting groups (City/Suburb and Town/Rural), identified possible future research avenues, but the groupings are somewhat artificial and may ignore real differences that exist between characteristics of schools in cities and suburbs or towns and rural settings. Finally, though this study focused on two contextual variables, there are a number of additional variables that may interact with quality of intervention implementation, and which may benefit from exploration in future studies.

Implications and Future Directions

This study investigated the effects of quality of PBIS implementation and school contextual variables on behavioral and academic outcomes. Of particular interest was whether there were any interactions between PBIS quality and school economic status. In examinations of suspension rates and percentages of students demonstrating proficiency in reading and math, no interactions were found. Not surprisingly, main effects were found of economic status on all dependent variables, with higher rates of suspension and lower levels of academic proficiency found in Low Economic schools. Main effects of PBIS Level were also found on math and reading proficiency. Schools that implemented PBIS with the highest level of quality had higher percentages of students who attained proficiency than schools that implemented with lower levels of quality. Surprisingly, quality of PBIS implementation was not found to have a significant effect on suspension rates. An

exploratory analysis that included PBIS Level, Economic Status, and Setting found a significant interaction between PBIS Level and Setting with regard to suspension. Significant main effects of PBIS Level as well as an interaction between Economic Status and Setting were found on reading and math proficiency. The findings of this study have several implications for research as well as for systems that are implementing models of prevention or attempting to identify methods for encouraging and recognizing high-quality implementation.

With regard to suspension rates, several patterns that were identified in the data may explain the lack of a significant relationship with quality of PBIS implementation. Moderate/High Economic Status schools had fairly low suspension rates at the lowest level of implementation, which limited the impact that increased quality of PBIS might be expected to have. Simply put, as a whole, schools at the lowest level of recognition appeared to be functioning fairly well with regard to application of exclusionary discipline. Future research regarding the effects of quality of PBIS implementation in these schools could investigate other behavioral indicators, such as students' perceptions of school or the incidence of bullying.

Low Economic Status schools, on the other hand, had a wide range of suspension rates. The schools with the highest suspension rates were in PBIS Level 2. With the assumption that those schools did not differ from the rest of the population of schools in some unidentified way, PBIS implementation was not effectively addressing those outlier schools' needs. Further research is needed to investigate the effects of quality of PBIS among Low Economic Status schools. In order to clarify this relationship, it may be helpful to include a comparison group of schools not implementing PBIS. Additionally, it may be

helpful to examine the suspension rates at Low Economic Status schools as they implement PBIS over time. Finally, as noted above, studying other outcomes in addition to suspension rates, such as attendance, office discipline referrals, or students' perceptions, may provide useful information.

The exploratory analysis suggests that continued research into possible interaction effects of PBIS quality and contextual variables is needed. Though the sample size was small, differences were seen between schools grouped by PBIS level, economic status, and setting. In particular, the interaction between PBIS Level and Setting was interesting. The data appeared to suggest better behavioral outcomes at higher levels of PBIS quality within City/Suburb schools, but not within Town/Rural schools. Additional research with a larger number of schools is needed in order to provide more information and to increase the likelihood of identifying significant relationships between PBIS quality, contextual variables, and behavioral and academic outcomes.

This study may also have practical implications for the PBIS recognition system being used in North Carolina and for similar systems used to promote high-quality implementation of prevention frameworks. The goal of the recognition system was to identify critical components necessary for quality implementation, and to recognize schools that achieved increased levels of quality. Though average suspension rates were lower among schools earning the highest level of PBIS recognition than among other levels, that difference was not significant. The study identified significant main effects of PBIS Level on reading and math achievement, but additional research is needed to identify what aspects of quality may be tied to higher achievement.

Among the three levels of PBIS quality, there were differences in the extent of training and fidelity of implementation. Additionally, schools earning the highest level of recognition identified positive trends in academic and behavioral data. Research that explores each of these variables (level of training, fidelity, and data-based decision making) may help to identify whether a given variable is more closely related to improved outcomes than other variables. This research could be used to fine-tune the criteria for recognition. In North Carolina, a prevention model that incorporates both academic and behavioral support, called Multi-Tiered Systems of Support (MTSS), is being introduced. This study indicates that a similar recognition system may be useful in promoting high-quality implementation of the MTSS framework, though additional research is needed to examine what elements are most critical, and whether quality of implementation interacts with contextual variables.

REFERENCES

- Algozzine, B., Wang, C., & Violette, A. S. (2010). Reexamining the relationship between academic achievement and social behavior. *Journal of Positive Behavior Interventions*.
- American Psychological Association Zero Tolerance Task Force. (2008). Are zero tolerance policies effective in the schools?: an evidentiary review and recommendations. *The American Psychologist*, 63(9), 852.
- Anderman, E. M. (2002). School effects on psychological outcomes during adolescence. *Journal of Educational Psychology*, 94(4), 795.
- Barnes, J., Belsky, J., Broomfield, K. A., & Melhuish, E. (2006). Neighbourhood deprivation, school disorder and academic achievement in primary schools in deprived communities in England. *International Journal of Behavioral Development*, 30(2), 127-136.
- Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions*, 12(3), 133-148.
- Bradshaw, C. P., Sawyer, A. L., & O'Brennan, L. M. (2009). A social disorganization perspective on bullying-related attitudes and behaviors: The influence of school context. *American Journal of Community Psychology*, 43(3-4), 204-220.
- Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *The Journal of Educational Research*, 90(5), 269-277.
- Christie, C. A., Nelson, C. M., & Jolivette, K. (2004). School characteristics related to the use of suspension. *Education and Treatment of Children*, 509-526.
- Cohen, R., Kincaid, D., & Childs, K. E. (2007). Measuring school-wide positive behavior support implementation development and validation of the benchmarks of quality. *Journal of Positive Behavior Interventions*, 9(4), 203-213.
- Darney, D., Reinke, W. M., Herman, K. C., Stormont, M., & Ialongo, N. S. (2013). Children with co-occurring academic and behavior problems in first grade: Distal outcomes in twelfth grade. *Journal of school psychology*, 51(1), 117-128.
- Domitrovich, C. E., Bradshaw, C. P., Poduska, J. M., Hoagwood, K., Buckley, J. A., Olin, S., ... & Ialongo, N. S. (2008). Maximizing the implementation quality of evidence-based preventive interventions in schools: A conceptual framework. *Advances in School Mental Health Promotion*, 1(3), 6-28.

- Fidell, L. S., & Tabachnick, B. G. (2003). Preparatory data analysis. In J. A. Schinka & W. F. Velicer (Eds.), *Handbook of psychology, Volume 2: Research Methods in Psychology* (115-141). Hoboken, NJ: Wiley and Sons.
- Forsyth, C. J., Howat, H., Pei, L. K., Forsyth, Y. A., Asmus, G., & Stokes, B. R. (2013). Examining the infractions causing higher rates of suspensions and expulsions: Racial and ethnic considerations. *Laws*, 2(1), 20-32.
- Gietz, C., & McIntosh, K. (2014). Relations between student perceptions of their school environment and academic achievement. *Canadian Journal of School Psychology*, 0829573514540415.
- Glew, G. M., Fan, M. Y., Katon, W., Rivara, F. P., & Kernic, M. A. (2005). Bullying, psychosocial adjustment, and academic performance in elementary school. *Archives of pediatrics & adolescent medicine*, 159(11), 1026-1031.
- Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The Achievement Gap and the Discipline Gap Two Sides of the Same Coin?. *Educational Researcher*, 39(1), 59-68.
- Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure?. *Child development*, 76(5), 949-967.
- Heck, R. H., & Mayor, R. A. (1993). School characteristics, school academic indicators and student outcomes: Implications for policies to improve schools. *Journal of Education Policy*, 8(2), 143-154.
- Hemphill, S. A., Toumbourou, J. W., Smith, R., Kendall, G. E., Rowland, B., Freiberg, K., & Williams, J. W. (2010). Are rates of school suspension higher in socially disadvantaged neighbourhoods? An Australian study. *Health Promotion Journal of Australia*, 21(1), 12-18.
- Hoagwood, K., & Johnson, J. (2003). School psychology: A public health framework: I. From evidence-based practices to evidence-based policies. *Journal of School Psychology*, 41(1), 3-21.
- Hope, T. L., & Bierman, K. L. (1998). Patterns of home and school behavior problems in rural and urban settings. *Journal of school psychology*, 36(1), 45-58.
- Horner, R. H., Sugai, G., Smolkowski, K., Eber, L., Nakasato, J., Todd, A. W., & Esperanza, J. (2009). A randomized, wait-list controlled effectiveness trial assessing school-wide positive behavior support in elementary schools. *Journal of Positive Behavior Interventions*.
- Horner, R. H., Todd, A. W., Lewis-Palmer, T., Irvin, L. K., Sugai, G., & Boland, J. B. (2004). The School-Wide Evaluation Tool (SET) A Research Instrument for Assessing School-Wide Positive Behavior Support. *Journal of Positive Behavior Interventions*, 6(1), 3-12.

- Hughes, C.D. & Algozzine, B. (2008). *Implementation Inventory Online*. Retrieved from: <https://r7behaviorsupport.wikispaces.com/file/view/IIO+Doc.doc>
- Irwin, D., & Algozzine, B. (2005). *North Carolina Positive Behavior Support Initiative Evaluation Report*. Raleigh, NC: Department of Public Instruction, Exceptional Children Division, Behavioral Support Services
- Irwin, D., & Algozzine, B. (2006). *North Carolina Positive Behavior Intervention & Support Initiative Evaluation Report 2004-2005*. Raleigh, NC: Department of Public Instruction, Exceptional Children Division, Behavioral Support Services
- Irwin, D., & Algozzine, B. (2007). *North Carolina Positive Behavior Intervention & Support Initiative Evaluation Report 2005-2006*. Raleigh, NC: Department of Public Instruction, Exceptional Children Division, Behavioral Support Services
- Irwin, D., & Algozzine, B. (2008). *North Carolina Positive Behavior Intervention & Support Initiative Evaluation Report 2006-2007*. Raleigh, NC: Department of Public Instruction, Exceptional Children Division, Behavioral Support Services
- Jargowsky, P. A., & El Komi, M. (2011). Before or after the bell? School context and neighborhood effects on student achievement. In H. B. Newburger, E. L. Birch, & S. M. Wachter (Eds.), *Neighborhood and life chances: how place matters in modern America* (50-72). Philadelphia, PA: University of Pennsylvania Press.
- Lassen, S. R., Steele, M. M., & Sailor, W. (2006). The relationship of school-wide positive behavior support to academic achievement in an urban middle school. *Psychology in the Schools*, 43(6), 701-712.
- Luiselli, J. K., Putnam, R. F., Handler, M. W., & Feinberg, A. B. (2005). Whole-school positive behaviour support: effects on student discipline problems and academic performance. *Educational Psychology*, 25(2-3), 183-198.
- Ma, X., & Klinger, D. A. (2000). Hierarchical linear modelling of student and school effects on academic achievement. *Canadian Journal of Education/Revue canadienne de l'education*, 41-55.
- McDougal, J. L., Sheila, M. C., & Martens, B. K. (2000). Using organizational change procedures to promote the acceptability of prereferral intervention services: The school-based intervention team project. *School Psychology Quarterly*, 15(2), 149.
- McIntosh, K., Bennett, J. L., & Price, K. (2011). Evaluation of social and academic effects of school-wide positive behaviour support in a Canadian school district. *Exceptionality Education International*, 21(1), 46-60.
- Mendez, L. M. R., Knoff, H. M., & Ferron, J. M. (2002). School demographic variables and out-of-school suspension rates: A quantitative and qualitative analysis of a large, ethnically diverse school district. *Psychology in the Schools*, 39(3), 259-277.

- Miles, S. B., & Stipek, D. (2006). Contemporaneous and longitudinal associations between social behavior and literacy achievement in a sample of low-income elementary school children. *Child development*, 77(1), 103-117.
- Morrison, G. M., & D'Incau, B. (1997). The web of zero-tolerance: Characteristics of students who are recommended for expulsion from school. *Education and Treatment of Children*, 316-335.
- Muscott, H. S., Mann, E. L., & LeBrun, M. R. (2008). Positive Behavioral Interventions and Supports in New Hampshire Effects of Large-Scale Implementation of Schoolwide Positive Behavior Support on Student Discipline and Academic Achievement. *Journal of Positive Behavior Interventions*, 10(3), 190-205.
- National Center for Education Statistics. (n.d.). *Identification of rural locales*. Retrieved from https://nces.ed.gov/ccd/rural_locales.asp
- NCDPI/North Carolina Testing Program. (2014). *Technical brief-EOG and EOC assessments*. Retrieved from <http://www.ncpublicschools.org/docs/accountability/testing/eoceogtechbrief14.pdf>
- Noltemeyer, A., & Mcloughlin, C. S. (2010). Patterns of Exclusionary Discipline by School Typology, Ethnicity, and Their Interaction. *Penn GSE Perspectives on Urban Education*, 7(1), 27-40.
- North Carolina Department of Public Instruction. (n.d.). *Free and reduced student data by site: Public school year-to-date data 2013-2014*. Retrieved from <http://www.ncpublicschools.org/docs/fbs/resources/data/freereduced/2013-14freereduced.xls>
- North Carolina Department of Public Instruction. (2014). *Reliability of the North Carolina End-of-Grade and End-of-Course assessments*. Retrieved from <http://www.ncpublicschools.org/docs/accountability/testing/eogeocreliabilities14.pdf>
- Pas, E. T., & Bradshaw, C. P. (2012). Examining the association between implementation and outcomes. *The journal of behavioral health services & research*, 39(4), 417-433.
- Pas, E. T., Waasdorp, T. E., & Bradshaw, C. P. (2014). Examining Contextual Influences on Classroom-Based Implementation of Positive Behavior Support Strategies: Findings from a Randomized Controlled Effectiveness Trial. *Prevention Science*, 1-11.
- Payne, A. A., & Eckert, R. (2010). The relative importance of provider, program, school, and community predictors of the implementation quality of school-based prevention programs. *Prevention Science*, 11(2), 126-141.
- Payne, A. A., Gottfredson, D. C., & Gottfredson, G. D. (2006). School predictors of the intensity of implementation of school-based prevention programs: Results from a national study. *Prevention Science*, 7(2), 225-237.

- PBIS Exemplar schools 2013-2014*. (n.d.). Retrieved from:
<http://behaviorsupport.ncdpi.wikispaces.net/file/view/Exemplar%20Schools%202014.pdf>
- PBIS Green Ribbon schools 2013-2014*. (n.d.). Retrieved from:
<http://behaviorsupport.ncdpi.wikispaces.net/file/view/Green%20Ribbon%20Schools%202014.pdf/>
- PBIS Model schools 2013-2014*. (n.d.). Retrieved from:
<http://behaviorsupport.ncdpi.wikispaces.net/file/view/Model+Schools+2014.pdf>
- Public Schools of North Carolina. (n.d.1). *PBIS school recognition criteria*. Retrieved from:
<http://www.dpi.state.nc.us/docs/positivebehavior/data/recognition/pbsschoolrecognition.pdf>
- Public Schools of North Carolina. (n.d.2). *PBIS Team Training Module 1: Universal Implementation*. Retrieved from:
<http://www.dpi.state.nc.us/docs/positivebehavior/implementation/modules/module1/slides.ppt>
- Public Schools of North Carolina. (2013). *North Carolina Positive Behavior Intervention and Support 11-12 evaluation report*. Retrieved from:
<http://www.ncpublicschools.org/docs/positivebehavior/data/evaluation/2011-12/2012-eval-report.pdf>
- Public Schools of North Carolina. (2014). *North Carolina Positive Behavior Intervention and Support 12-13 Evaluation Report*. Retrieved from:
http://pbis.ncdpi.wikispaces.net/file/view/14pbisreport_FINAL.pdf/526770366/14pbisreport_FINAL.pdf
- Reinke, W. M., Herman, K. C., & Stormont, M. (2013). Classroom-level positive behavior supports in schools implementing SW-PBIS identifying areas for enhancement. *Journal of Positive Behavior Interventions*, 15(1), 39-50.
- Reynolds, H., Irwin, D., & Algozzine, B. (2009). *North Carolina Positive Behavior Intervention and Support initiative evaluation report 2007-2008*. Raleigh, NC: Department of Public Instruction, Exceptional Children Division, Behavioral Support Services.
- Rohrbach, L. A., Grana, R., Sussman, S., & Valente, T. W. (2006). Type II translation transporting prevention interventions from research to real-world settings. *Evaluation & the Health Professions*, 29(3), 302-333.
- Simonsen, B., Eber, L., Black, A. C., Sugai, G., Lewandowski, H., Sims, B., & Myers, D. (2012). Illinois Statewide Positive Behavioral Interventions and Supports Evolution and Impact on Student Outcomes Across Years. *Journal of Positive Behavior Interventions*, 14(1), 5-16.

- Sirard, J. R., Ainsworth, B. E., McIver, K. L., & Pate, R. R. (2005). Prevalence of active commuting at urban and suburban elementary schools in Columbia, SC. *American Journal of Public Health, 95*(2), 236.
- Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of educational research, 75*(3), 417-453.
- Skiba, R.J., & Peterson, R.L. (1999). The dark side of zero tolerance: Can punishment lead to safe schools? *Phi Delta Kappan, 80*, 372–382.
- Skiba, R. J., & Peterson, R. L. (2000). School discipline at a crossroads: From zero tolerance to early response. *Exceptional Children, 66*(3), 335-346.
- Stewart, E. A. (2003). School social bonds, school climate, and school misbehavior: A multilevel analysis. *Justice Quarterly, 20*(3), 575-604.
- Sugai, G., & Horner, R.R. (2002). The evolution of discipline practices: School-Wide Positive Behavior Support. *Child & Family Behavior Therapy, 24*, 23-50.
- Sugai, G., & Horner, R. R. (2006). A promising approach for expanding and sustaining school-wide positive behavior support. *School psychology review, 35*(2), 245.
- Sugai, G., Horner, R. H., Dunlap, G., Hieneman, M., Lewis, T. J., Nelson, C. M., ... & Ruef, M. (2000). Applying positive behavior support and functional behavioral assessment in schools. *Journal of Positive Behavior Interventions, 2*(3), 131-143.
- Sugai, G., & Simonsen, B. (2012). Positive behavioral interventions and supports: History, defining features, and misconceptions. *PBIS Revisited, 1*(1).
- Technical Assistance Center on Positive Behavioral Interventions and Supports. (2010). *Implementation blueprint and self-assessment: Positive Behavioral Interventions and Supports*. Retrieved from: <https://www.pbis.org/blueprint/implementation-blueprint>
- Ware, W. B. (2014). *The general linear model: Multiple regression and analysis of variance*. Chapel Hill, NC: School of Education, University of North Carolina at Chapel Hill.